Functional Attributes and Self-Assembly Behavior of a Shell Protein from the 1,2-Propanediol Utilization Prokaryotic Metabolosome

Gaurav Kumar

A thesis submitted for the partial fulfillment of the degree of Doctor of Philosophy

Institute of Nanoscience and Technology (INST) Mohali
Knowledge City, Sector 81, SAS Nagar, Manauli, P.O. 140306
Punjab, India.

Indian Institute of Science Education and Research (IISER) Mohali
Knowledge City, Sector 81, SAS Nagar, Manauli, P.O. 140306
Punjab, India.

February 2023

Dedicated to my beloved grandparents...

Declaration

The work presented in this thesis has been carried out by me under the guidance of Dr. Sharmistha Sinha at the Institute of Nano Science and Technology, 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.

GAURAV KUMAR

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. SHARMISTHA SINHA

Acknowledgements

As I present my work compiled in this thesis, I must admit that this would not have been possible without support, help, guidance and love I have received from so many people. Right from my schooling days to research in academics, a lot of people have made indubitable contributions towards building my career in research. I will therefore take the opportunity to express my sincere gratitude to everyone who supported me throughout and all those who made significant impact in my life.

It has been a wonderful journey filled with many ups and downs, sporadic moments of pleasure and pain. It's never easy to venture onto a journey to study the molecular details of a complex system like microcompartment. You require the guidance and support of a supervisor who teaches you and gives you the freedom to think, reason as well as self-critic. I was fortunate enough to have **Dr. Sharmistha Sinha** as my mentor and guide in this journey. Her experience in protein biophysics and imagination skills really fascinated me and helped me to delve into the complex organization of microcompartments. Her ability to critic and listen to counter arguments made scientific discussions fun and enabled me to formulate a few intriguing questions and hypothesis. I will always cherish the brainstorming sessions we had together while designing experiments and writing projects. She always encouraged me to participate in project writing, talks and conferences. This made me more confident and independent and improved my ability to tackle challenging experiments. It also strengthened my love for science communication and writing. Her sustained encouragement during successful experiments or during failures motivated me to keep going. Her dedication towards science and management skills makes her an efficient team leader. I hope someday I will be able to inculcate her multi-tasking skills in myself. I am confident that the knowledge and training I have acquired being her Ph.D. student will certainly help in shaping my future research accomplishments.

I would like to thank my thesis committee members, **Dr. Sangita Roy, Dr. Asish Pal** and **Dr. Surajit Karmakar** for their time, keen interest in my work and for providing thoughtful suggestions. Their critical comments helped me to improve the quality of my experimental designs and data analysis.

I express my sincere gratitude to **Dr. Sabyasachi Rakshit**, IISER Mohali for being kind and very supportive. I thank him for allowing me to work in his lab during the first three years of my Ph.D. I enjoyed having scientific discussions with him for our collaborative projects. His

with you on our collaborative project and I am happy that we could produce some fantastic results. You constantly provided me with new ways to see experiments, which broadened my thinking. Your criticism of my experiments and storyline improved the quality of my thesis. You have helped me out as a friend when I most needed it. Without you, it would have been very difficult for me to make it through the final six months of my tenure. I thoroughly enjoyed and will always remember our discussions about science and life over a cup of cappuccino. Also, **Toto** will always remain close to my heart.

Simerpreet, you were the most experienced one in the lab and I guess it made me feel more positive every time. You taught me cloning techniques and gave me hands on experience with molecular biology. For this, I will always be grateful to you. I always saw a fighter in you and I believe you have immense potential to achieve excellence. Although we never collaborated on any project I hope in later years we get a opportunity to work together.

As days passed, our lab grew in size and we were joined by some fun-loving and energetic individuals. **Dimple**, you are someone who is always eager to help. I appreciate the fact that you always take the initiative in every lab activity, even though it requires a lot of effort. This is truly a beautiful character, and I hope to cultivate some of it in myself. **Preeti**, you are the reason why we never have to worry about things going misplaced. Your commitment to maintaining a clean and organized lab is unparalleled. Your management skills, outsourcing and getting things done are indeed fascinating and I hope to nurture these characters in myself. I enjoyed my time with you here. Our tea-time talks and discussions along with **Bansi** will always remain special.

Harpreet, you are weirdly entertaining and I have always felt relaxed listening to you and laughing around you without any hesitation. **Rose** and **Aarcha** are new members of our lab. **Rose** is quite interesting and well-read person. If I ever needed to discuss any chemistry related stuffs, I felt more comfortable in asking **Rose** than googling. **Aarcha** is very patient in her work. Whether her experiments are a success or a failure, her ability to maintain consistency is noteworthy.

During my Ph.D. I had the opportunity to learn a lot from SR Lab members. **Jagadish,** I thank you for introducing me to TIRF, TCSPC and MD simulation. I am happy that we collaborated and generated some very interesting data. **Surbhi, Nisha** and **Sai**, I thank all of you for your support and help during my stay in SR Lab. **Sayan**, you were here for a short duration but I thoroughly enjoyed your company. You helped me in troubleshooting cloning protocols during the initial days of my Ph.D. I will always remember our frequent coffee breaks, chit-chatting and gossiping about random people.

A few teachers who had a significant impact on my life deserve my gratitude. I thank **Dr. Amitava Mukherjee**, VIT University for giving me the opportunity to work independently in his lab during my masters. His guidance and teachings helped me to become better as a researcher. I am grateful to **Prof. B.B. Sahay** for teaching me physics as well as teaching me how to imagine and visualize everything in terms of figures and models. His teachings have played a significant role during my Ph.D. while designing experiments and postulating hypothesis. I thank **Sir. Bhawender** for helping me learn mathematical concepts in simplest possible way. I also wish to thank **Sir Tapan Mukherjee** for the fascinating science classes in school. I must admit that **Sir Tapan** was responsible for my growing appreciation of science over other subjects.

All these years, I was fortunate enough to be motivated, supported and loved by beautiful group of friends. **Akash, Rahul** and **Tanaya,** you were and will always be another family away from my family. I have learned so much from you guys. You accepted me as I was and inspired me to become better. We had our moments of joy and sorrow. **Akash** and **Rahul**, you were there when I felt confident, hopeful and positive. You did not leave me when I broke down and my self-confidence hit rock bottom. Thank you for everything. **Tanaya**, your goodness, kindness and selfless care is what makes you a great friend. I will always remember you calling me a night before exams for para phrasing topics for the following day. Your determination to persevere despite all odds makes you an inspiration.

Family is always there to support you, no matter where you are or what you're doing. Because of everything my family has done for me, I want to express my genuine gratitude to them. I have watched **my father** work incredibly hard throughout his life to provide for our family and for our education. He made every effort to guarantee that we would never have to compromise the quality of our education. I'll never forget how, when I was a child, he would take a hibiscus flower (china rose) and slice it open with a blade to show me the stamen and pollen sac. Every evening, I would ask him to read to me from the school science textbook. These little acts deepened my interest in science and research. **My mother,** it would take an entire lifetime to thank you for what you have done for me and my sister. You fought for our education and good schooling. Thank you for understanding and supporting me in every decision I took. I know nothing matters to you more than seeing us around you. Thank you for understanding when I couldn't be there. Thank you for understanding when I didn't come home. My dear sister **Shreya**, you are truly an amazing human being. You have triumphed over all obstacles that stood in your way and have overcome all failures, struggles and

hurdles. I'm privileged to have you as a sister. I thank you for being there for me, mom and dad.

I would like to thank **my grandparents**, who loved me whole-heartedly and always encouraged me to excel in life. I thank **Meso** for fostering an environment of reading and learning at home. They are all greatly missed.

I would also like to thank all my cousin brothers and sisters, **Sweety, Cuty, Jishu, Mishu** for sweet childhood memories we share. **Vishwajeet** and **Abhijeet,** I am grateful to you for being such wonderful brothers. You have taught me to be tough and strong and have inspired me to face real life challenges. **Abhishek (Nikki da),** I thank you for supporting me all along to pursue a career in science research. I am fortunate to have an intellectual elder brother like you. Thank you **Shikha** for your love and support you have shown for our family. I thank all the members of my family, **Masi, Meso, Kaka, Kaki, Choto Dadu and Dida, Pise, Pisi, Topi Mama and Mami, Mitu, Itu** for their constant support and encouragement. Thank you **Sonali** for everything you have done for our family.

I would like to thank my colleagues in INST, **Ashish, Ayushi, Tanmay, Gurpreet, Avinash, Mujeeb, Akriti, Hari** for helping me out in my research work. **Mahima**, I thank you for always being prompt to co-host events and conferences. It was fun doing stage work with you.

Himadri, thank you for being so kind and loving. Thank you for everything you have done for me. It means a lot. I will never forget you morning breakfasts. **Shobhit**, you are really a great guy and a wonderful friend. You have taught me how to take responsibility. I have always liked our intellectual discussions and conversations. I will always admire your attitude of generosity and compassion.

Without the assistance and support of everyone named here, this endeavor would not have been possible, and I will always be grateful for what they did.

Gaurav Kumar

Abbreviations

MCP	**Microcompartment**
BMC	**Bacterial Microcompartment**
PDU	**Propanediol utilization**
PD	**Propanediol**
MBTH	**3-Methyl-2-benzothiazoline**
BPER-II	**Bacterial Protein Extraction Reagent**
CBC	**Calvin Benson Cycle**
CA	**Carbonic anhydrase**
DSF	**Differential scanning fluorimetry**
DDH	**Diol dehydratase**
DLS	**Dynamic Light scattering**
EUT	**Ethanolamine UtilizationMicrocompartment**
IPTG	**Isopropyl β-d-1-thiogalactopyranoside**
PMSF	**Phenylmethane sulfonyl fluoride**
NHS	**N-Hydroxy succinamide**
NAD	**Nicotinamide Adenine Dinucleotide**
PEG	**Polyethylene glycol**
PAGE	**Polyacrylamide Gel Electrophoresis**
RuBisCo	**Ribulose-1,5-bisphosphate carboxylase/oxygenase**
SDS	**Sodium dodecyl sulphate**
TSA	**Thermal shift assay**
TEM	**Transmission electron Microscopy**
UV	**Ultraviolet spectroscopy**
Vit B$_{12}$	**Vitamin B12**
MD	**Molecular Dynamics**

Abstract

Complex subcellular organization enable cells to synergize multiple metabolic reactions simultaneously. Many species of bacteria have the ability to compartmentalize enzymatic cascade within a nano-sized protein compartments called bacterial microcompartments. These microcompartments have evolved to perform catabolic or anabolic reactions and some of them aid bacteria to survive under energy dearth environment. Within my research, I have focused on deciphering the assembly and function of the proteins that make up the architecture of a bacterial microcompartment involved in 1,2-propanediol metabolism (1,2-propanediol utilization microcompartment) in *Salmonella enterica* LT2. The component proteins of these microcompartments are encoded by genes belonging to a single operon. The operon encodes for two types of proteins, the shell proteins that form the outer cover of the compartment and the enzymes that are encapsulated within and work in cascade. My study shows that a major shell protein called PduBB', exhibits high thermal stability and preserves the catalytic activity and native conformation of an encapsulated diol dehydratase enzyme called PduCDE. The individual components of the shell protein (PduB and PduB') in isolation fail to influence the activity of PduCDE and do not show any protective role. The combination of PduB and PduB' provides solubility and stability to the shell protein PduBB', improving its association with the enzyme PduCDE. The longer component PduB has extra 37 amino acids N-terminal extension that has a short helical segment followed by a disordered region. Similarly, the middle subunit (subunit D) of PduCDE also has a helical portion followed by disordered region. A combination of biochemical, spectroscopic and computational modeling studies suggests that the N-terminal extensions of PduB and PduD might be involved in mediating shell protein-enzyme interaction. We propose that the disordered regions provide flexibility to the N-terminal extensions of shell protein and enzyme, facilitating their association mediated by the helical segments. PduBB' displays an interesting self-assembly behavior and under crowded environment and appropriate ionic strength, it undergoes liquid-liquid phase separation. While PduB' has a high self-associating property displaying liquid-solid transition, PduB has the lowest tendency to undergo phase separation. This implies that the combination of two proteins in PduBB' results in a well-balanced self-assembly behavior. The co-phase separation of shell protein PduBB' and enzyme PduCDE enhances the catalytic efficiency of the enzyme, highlighting the significance of shell-enzyme association and phase separation in improving the catalytic

performance of the enzyme. My study has two major implications. First, it suggests that formation of PduMCP is likely to be mediated by a combination of protein-protein interaction and phase separation, where disordered N-terminal regions provide flexible association between shell protein and enzyme. Second, the chaperone like behavior of a major shell protein of contributes towards higher thermal stability and catalytic activity of an encapsulated native enzyme.

SYNOPSIS

A remarkable feature of living cells is that they compartmentalize biological macromolecules inside a limited volume and space. As a result of molecular confinement, molecules can stay close together and collaborate to enhance biological processes. Subcellular compartmentalization results in the formation of organelles that separate biomolecules like proteins, carbohydrates, and nucleic acids from the surrounding cellular environment. Notably organelles are not exclusive to eukaryotes but are also produced by wide number of prokaryotes. Many bacterial species harbor tiny compartment like bodies made of proteins called bacterial microcompartments (MCP). These MCPs are polyhedral in shape and appear in the size range of 100-200 nm. They are exclusively made of proteins and are devoid of nucleic acids (DNA/RNA) or outer lipid membrane. They are made of thousands of protein subunits that are encoded by genes usually belonging to one single operon. The outer cover of the MCPs are made of shell proteins and the enzymes are encapsulated within. The shell proteins act as selectively permeable layer for the transport of substrates and other metabolites. The complex nature and composition of MCPs make they an interesting subject of study in the context of structure function relations and self-assembly of proteins.

The present thesis attempts to work on the fundamental questions in compartment biology. Why is compartmentalization favored by cells? What factors trigger the proteins to self-assemble into compartment like structures? How do the structural features of the component proteins impact their functions in the compartmentalization and metabolism? To address these questions, I have used 1,2-propanediol utilization compartments (PduMCP) that are produced by *Salmonella enterica* LT2 under energy dearth environment. PduMCPs enable the bacteria to utilize 1,2-propandiol (1,2-PD) as an alternate carbon source when there is a scarcity of food. It is made of 8 different shell proteins (PduA, B/B', J, M, N, K, T and U) and rest are the enzymes that participate in 1,2-PD metabolism, NADH/MADH recycling, and activation of cofactor cyanocobalamin (Vit-B12). Among the shell proteins, PduBB' is quite interesting as it is made of two proteins PduB and PduB', both encoded by tow alternate start codons in single gene. The longer form of the shell protein (PduB) has extra 37 amino acids N-terminal extension. The signature enzyme of the PduMCP is diol dehydratase PduCDE. It is the first enzyme in the enzymatic cascade in PduMCp and has three subunits PduC, PduD and PduE. Among the three subunits the largest is PduC that is also the catalytic subunit of the PduCDE holoenzyme and binds to substrate 1,2-PD as well as cofactor Vit-B12.

In *Chapter 1*, I begin by describing the complex organization of cells and its role on the formation of organelles. I briefly explain how cells have evolved to produce both organelles with outer lipid barrier and membraneless organelles. MCPs are also devoid of outer lipid bilayer and the enzymatic core is covered by a blanket of shell proteins. Based on the reactions they perform, MCPs are categorized into carboxysomes and metabolosomes. Carboxysomes are found in photosynthetic bacteria and participate in CO2 fixation. While metabolosomes are catabolic reactors that catabolize alternate carbon based substrates, aiding bacteria to survive under food scarcity. I have given a detailed description supported by schematic representation on the functions and structural features of carboxysomes and metabolosomes. A comparison between the structure and composition of two different types of carboxysomes (α-carboxysome and ß-carboxysome) has also been tabulated. A description on the outer shell of MCPs provides information on the structure and assembly of the shell proteins, where I describe different BMC domain families of shell proteins. Using models of shell proteins, I explain how homomerization of shell protein subunits result in the formation of cyclic-hexamer cyclic pseudo-hexamer or pentamer. The cyclic-hexamers or pseudo-hexamers have central pores that are suggested to help in metabolite transport. This is followed by timeline of MCP research from 1950s-present, from the perspective of biophysics. In this section I provide a concise review on the contribution of electron microscopy, molecular biology, fluorescence microscopy and various analytical techniques in understanding the fundamentals of MCPs. Electron microscopy has enabled researches to visualize the size, morphology and shape of different MCPs and differentiating between the luminal arrangements of α-carboxysome and ß-carboxysome. Gene deletion studies has revealed the functions of various shell proteins and importance of their terminal regions. Recent studies using fluorescence microscopy has elucidated the dynamics of the shell proteins and enzymes of carboxysomes. These reports have paved the way for the development of MCP based synthetic reactors. Overall, this section provides a summary of significant work done towards understanding the stricture, size, composition, physical properties and biogenesis of MCPs.

In the later part of the chapter I present the research theme of the thesis. A brief introduction to PduMCP is given followed by detailed sketch about its structure and function. In the final section of the chapter, research questions have been described. The questions span over the ideas related to molecular confinement and enzyme function, functions of shell proteins, shell protein-enzyme interaction, self-assembly of shell protein and enzyme and potential role of structured and unstructured regions in the functioning of PduMCP proteins.

These questions have been addressed using a barrage of biophysical and analytical techniques. I have combined spectroscopy, molecular biology, bioinformatics/computational techniques to generate results. In *Chapter 2*, I have elaborated on the various techniques used in the thesis. I have also tabulated the materials, strains, and reagents used in the study.

In *chapter 3*, I address the first set of questions related to molecular confinement and enzyme functions. This chapter explores enzyme assays, fluorescence spectroscopy, interferometry studies and molecular modeling to study the effect of PduBB' on the stability and catalytic efficiency of PduCDE. Using enzyme assay and fluorescence spectroscopy, I demonstrate that the enzyme molecules have higher stability and efficiency when encapsulated within PduMCP. Further the role of major shell protein PduBB' and its impact on the functioning of the signature enzyme diol dehydratase PduCDE has been studied. The free enzyme PduCDE in aqueous solution is labile and tends to denature when subjected to heat stress. Around 45% decrease in the activity of free enzyme PduCDE is seen at 45°C, while intact PduMCP maintains its optimum diol dehydratase activity till 50°C, and reduction in the PduMCP activity is seen at and above 55°C. PduBB', the major component of the outer shell contributes to the increased catalytic activity of PduCDE. PduBB' also prevents the unfolding and aggregation of PduCDE at higher temperatures. Using bio-layer interferometry, fluorescence anisotropy and molecular docking, the interactions of the shell proteins PduBB', N-terminal truncated PduB' and PduB (generated as a single mutant PduBM38L) with PduCDE have been probed. All the three variants of the shell protein (PduBB'/PduB/PduB') tend to interact with the enzyme PduCDE with micro molar affinity *in vitro*, but only PduBB' influences its activity and stability, emphasizing on the significance of the unique combination of PduB and PduB' in PduMCP stability.

In **chapter 4** of the thesis, I attempt to address perplexing question in MCP biology, which is to understand the mechanism which governs the formation of these small yet complex assemblages of proteins. Since PduMCP is a made of multiple shell proteins and enzymes, I take a minimalistic approach and use the major shell protein PduBB' as a paradigm to probe self-assembly dynamics in solution. For this study, I have used fluorophore labeled PduBB' and visualized its dynamics in solution phase under varying ionic strength and molecular crowding. The results show that the surrounding environment modulates the self-assembly behavior of PduBB'. Appropriate ionic strength and macromolecular crowding bring about liquid-liquid phase separation of the shell protein. Under crowded environment, salts display

unique property to drive the formation of shell protein liquid condensates. By probing the surface charge and hydrophobic patches of the shell protein, I conclude that the surface charge masking and kosmotropic effect of ions could be the probable underlying mechanism behind the salt mediated phase separation of PduBB'. The co-phase separation of PduCDE with PduBB' enables it to outperform the enzyme in isolation. A combination of spectroscopic and biochemical techniques shows the relevance of ions, especially Mg^{2+} in providing stability to intact PduMCP *in vitro*. While the N-terminal truncated PduB' has a high self-associating property displaying liquid-solid transition, the longer version of the shell protein PduB has the lowest tendency to undergo phase separation. This indicates that the N-terminal extension of PduBB' helps in optimization of the self-assembly of the shell protein. The higher self-associating property of PduB' is balanced by PduB which has higher solubility due to N-terminal extension. Together this study suggests that self-assembly of shell-protein and enzyme is triggered by a combination of protein-protein interactions and phase separation.

In ***chapter 5***, I have explored the distribution and potential roles of disordered regions in the integral proteins of 1,2-propanediol utilization microcompartments. The rationale behind this investigation is the fact that disordered regions have been reported to be crucial for the self-assembly and formation of many membrane less organelles. In this chapter, I use bioinformatics tools to identify the probable disordered regions in the shell proteins and enzyme of PduMCP. I observe that the major shell proteins of PduMCP (PduA/J/B) as well as minor shell protein PduK have intrinsically disordered terminal regions. The vertex protein PduN has disordered regions distributed throughout its polypeptide backbone. The disordered regions provide structural flexibility to the shell proteins and shell protein-enzyme interaction. In case of shell protein PduBB', the N-terminal region enhances its solubility and mediates a flexible interaction with the disordered N-terminal region of the PduD of the enzyme PduCDE. The presence of N-terminal region of PduD is important for its association with the shell protein PduBB'. Together these findings throw light on the importance of disordered regions in the self-assembly, providing flexibility to shell protein and mediating its binding with a native enzyme.

Chapter 6 summarizes the key findings of the work presented in the thesis. The results described and discussed in chapters 3-5 highlight the importance of major shell protein PduBB' in functioning and stability of the signature enzyme. The binding of PduBB' with PduCDE conserves the native conformation of the enzyme under thermal stress. PduBB'

exhibits fluid-like behavior and undergoes liquid-liquid phase separation in a crowded environment with appropriate ionic strength. Our findings imply that the 37 amino acid N-terminal extension controls the phase separation and self-assembly behavior of PduBB' in solution. While N-terminal truncated PduB' has a high self-associating property displaying liquid-solid transition, PduB with the N-terminal region is more soluble has the least potential to undergo phase separation. The combination of the two shell proteins in PduBB' results in a well-balanced self-assembly behavior, producing liquid droplets up to a very low shell protein concentration. We further observe that the co-phase separation of shell protein PduBB' and enzyme PduCDE enhances the catalytic efficiency of the enzyme. Our findings shed light on the importance of disordered regions in shell protein-enzyme interaction and self-assembly, providing insight into the microcompartment's biogenesis.

Publications included in thesis

Kumar G, Bari NK, Hazra JP, Sinha S. A Major Shell Protein of 1, 2-Propanediol Utilization Microcompartment Conserves the Activity of Its Signature Enzyme at Higher Temperatures. *ChemBioChem*. 2022 May 4: e202100694.

Kumar G, Sinha S. Biophysical approaches to understand and re-purpose bacterial microcompartments. *Current Opinion in Microbiology*. 2021 Oct 1; 63:43-51.

Kumar G, Hazra JP, Sinha S. Disordered regions endow structural flexibility to shell proteins and function towards shell–enzyme interactions in 1, 2-propanediol utilization microcompartment. *Journal of Biomolecular Structure and Dynamics*. 2022 Oct 23:1-1.

Kumar G, Sinha S. Self-assembly of Shell Protein and Native Enzyme in a Crowded Environment Leads to Catalytically Active Phase Condensates. *Biochemical Journal*. 2023 Jan 23:BCJ20220551.

Publications not included in thesis

Bedi S, **Kumar G**, Rose SM, Rakshit S, Sinha S. Barrier-free liquid condensates of nanocatalysts as effective concentrators of catalysis. Chem Commun (Camb). 2022 Jul 12. doi: 10.1039/d2cc03111f.

Bari NK, Hazra JP, **Kumar G**, Kaur S, Sinha S. Probe into a multi-protein prokaryotic organelle using thermal scanning assay reveals distinct properties of the core and the shell. *Biochimica et Biophysica* Acta (BBA)-General Subjects. 2020 Oct 1;1864(10):129680.

Bari NK, **Kumar G**, Hazra JP, Kaur S, Sinha S. Functional protein shells fabricated from the self-assembling protein sheets of prokaryotic organelles. *Journal of Materials Chemistry B*. 2020;8(3):523-33.

Bari NK, **Kumar G**, Bhatt A, Hazra JP, Garg A, Ali ME, Sinha S. Nanoparticle fabrication on bacterial microcompartment surface for the development of hybrid enzyme-inorganic catalyst. *ACS Catalysis*. 2018 Jul 24;8(9):7742-8.

Bari NK, **Kumar G**, Sinha S. The Wrappers of the 1, 2-Propanediol Utilization Bacterial Microcompartments. *Biochemical and Biophysical Roles of Cell Surface Molecules*. 2018:333-44.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Bacterial Microcompartment

1.1 Cellular organization

The complexity of life has fascinated biologists for years. It has long captured their imagination and can be seen in every life form, from bacteria to large animals and plants. The basic building block of life, the cell, is a remarkably sophisticated system made up of several smaller components including proteins, nucleic acids, carbohydrates, and lipids. These components work together as a team to keep life going [1]. It is fascinating to consider that more than four billion years of evolution could have led to such level of cellular intricacy we see today. At some point in the past, smaller molecules might have come together to form simple self-copying functional systems. With time, these simpler systems would have combined to give rise to more complex ones. Evolution has favored complexity for one good reason, and that is efficiency. A complex system is unquestionably more effective than a simpler one because it enables several components to remain adjacent and function in a highly choreographed manner, maximizing output [2],[3]. The question of how complexity arises out of simplicity has motivated decades of research into the understanding of the organization of living systems at a molecular level.

A cell is programmed to confine many bio-macromolecules within a limited space and volume through sub-cellular compartmentalization. Compartmentalization and localization of various entities such as enzymes, carbohydrates, and nucleic acids result in the formation of organelles that enable cells to sequester metabolic or catabolic reactions [4]. An interesting feature of many of these organelles such as mitochondria, chloroplasts, and Golgi bodies is that they are bound by an outer membrane made of lipid bilayer [5]. The protein channels in the lipid bilayer facilitate passage of small molecules across their lumen and cytoplasm [6],[7]. Apart from membrane-bound organelles there are membrane-less organelles that are devoid of outer lipid bilayer [8]. These include nucleolus, Cajal bodies and p-granules. Membraneless organelles have diverse roles in cells like cellular stress response [9], post translational modification [10], nucleic acid metabolism, RNA splicing and storage [11]. Last one decade has witnessed an accumulation of significant amount of information on the composition, structure and biogenesis of these membraneless organelles. Subcellular lipid free organelles are not just limited to eukaryotic cells; but have been discovered in many bacterial species as well, and one of them served as the subject of my Ph.D. research.

1.2 Bacterial microcompartments: complex yet beautiful

My introduction to bacterial microcompartments piqued my interest in subcellular organization. It was astounding to know that even bacteria with sizes between 1-2 microns could harbor tiny polyhedral compartment-like bodies. Bacterial microcompartments are functionally similar to eukaryotic organelles that encase multiple enzyme molecules within the core and carry our metabolic/catabolic reactions. Although they are devoid of lipid membrane, the enzymatic core is surrounded by semi permeable outer shell made of proteins. The protein shell layer acts as a barrier between the microcompartment lumen and bacterial cytoplasm and allow restricted movements of substrates, cofactors and products. These compartments are made of thousands of protein subunits and in many cases the constituent proteins are encoded by genes found in one large operon. Compartmentalization of the proteins within microcompartments helps in improving the enzymatic flux that may aid bacteria to survive better. The outer shell layer also acts as a diffusion barrier and prevents toxic/volatile intermediates from escaping out in to cytoplasm. Instead, the toxic intermediates are retained within the microcompartment lumen and are converted by downstream enzymes into non-toxic products. With the advancement in the field of metagenomics and incremental progress in the sequencing of bacterial genome, researchers have been able to able to detect genes in 23 known bacterial phyla. Sophisticated imaging techniques and crystallography have enabled us to delve into the structural details of the microcompartments [12] [13],[14],[15],[16] .

1.3 The outer shell of microcompartment

The outer shell layer of microcompartment is comprised of shell proteins that belong to one the following three categories: BMC-H, BMC-T and BMC-P [17] [18]. Among the three, BMC-H shell proteins are more abundant in microcompartments it contains BMC-H domain protein (Pfam00936) (**Fig. 1.1a**). Six BMC-H subunits self-assemble to form a cyclic hexamer with distinct concave and convex side. The hexamers further undergo later interactions through ionic interactions to form extended sheets. The sheet forming property of the shell protein is crucial as it allows the outer shell to blanket around the enzymatic core. BMC-T consists of two circularly permuted BMC-H domains that are fused in tandem (**Fig. 1.1b**). Three BMC-T subunits self-assemble to form trimer that structurally resemble BMC-H hexamer and hence they are called cyclic pseudohexamers. Similar to the BMC-H proteins, BMC-T proteins also exhibit sheet forming property. The central pore of both BMC-H

hexamer and BMC-T pseudohexamer can act as conduit for the transport of metabolites across the shell. BMC-P proteins are made of BMC-P domain (Pfam03319) in which five BMC-P domains self-assemble to form cyclic pentamer **(Fig. 1.1c)**. While BMC-H and BMC-T proteins have been suggested to form the sides or facets of the microcompartments, BMC-P proteins are suggested to occupy their vertices.
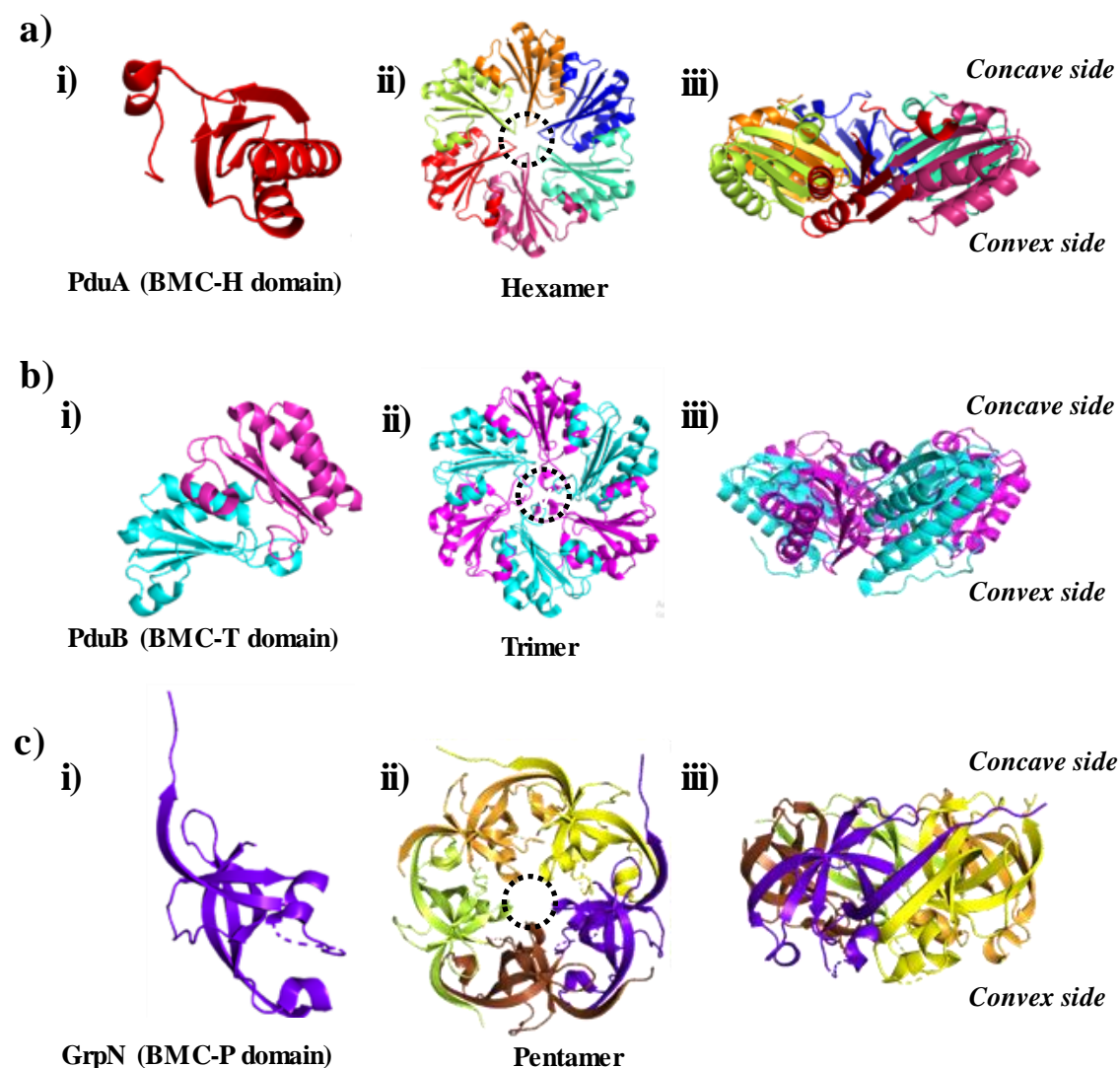
Fig. 1.1: a) Crystal structure of PduA (PDB: 3NGK), an example of BMC-H domain protein (i), Hexameric assembly of PduA (ii), Side view of PduA hexamer (iii) showing distinct concave and convex faces; b) Crystal structure of PduB (PDB: 4FAY), an example of BMC-T domain protein (i), Trimeric assembly of PduB (ii), Side view of PduB trimer (iii) showing distinct concave and convex faces; c) Crystal structure of GrpN (PDB: 4I7A), an example of BMC-P domain protein (i), Pentameric assembly of GrpN (ii), Side view of GrpN pentamer (iii) showing distinct concave and convex faces.

1.4 Encapsulated enzymes and functions

The widespread presence of BMC domain proteins across several bacterial species suggests that genes encoding BMC domain proteins have undergone horizontal gene transfer between species. A recent comprehensive genome analysis among 646 individual genome provided evidence in support of horizontal gene transfer idea and suggested the presence of a variety of microcoartments with diverse functions among 150 individual genomes. Based on the encapsulated pathways and reactions they carry out microcompartments have been grouped into carboxysomes and metabolosomes.

1.4.1 Carboxysomes

Carboxysomes encapsulate Ribulose-1,5-bisphosphate carboxylase/oxygenase or Rubisco are found in photosynthetic bacteria and many chemoautotrophs. The enzyme Rubisco is one of the most abundant enzyme on planet earth and is responsible for global CO_2 fixation [19],[20]. It participates in the Calvin-Benson cycle which produces glucose necessary for the growth of cells. The evolution of Rubisco dates back to the early earth period when it contributed significantly to lowering the CO_2 level in the earth's atmosphere. Surprisingly, its catalytic turnover has remained fairly low and it has a poor ability to differentiate between O_2 and CO_2. Non-specific binding of O_2 with Rubisco produces a toxic byproduct that may have a detrimental effect on the cells. Photosynthetic bacteria like cyanobacteria have evolved to overcome this problem by using their carbon concentrating mechanism or CCM. This mechanism involves the entrapment of atmospheric CO2 and converting them to HCO3- and feeding the HCO3- to the carboxysomes **(Fig. 1.2a)**. The carboxysomes contain encapsulated carbonic anhydrase enzyme that concerts HCO3- into CO2 which is then taken up by Rubisco present in the lumen of carboxysome. The outer shell layer of carboxysomes prevents the entry of O2 into the lumen of the carboxysome, thereby increasing the catalytic efficiency of the enzyme. Carboxysomes have been categorized into α-carboxysome and ß-carboxysome [20]. α-carboxysomes are present in cyanobacteria which are mostly found in the marine ecosystem and contain Form 1A type Rubisco. While ß-carboxysomes contain Form 1B Rubisco and are found in freshwater cyanobacteria. **Table 1.1** summarizes the differences between the two forms of carboxysomes based on existing reports in the literature.
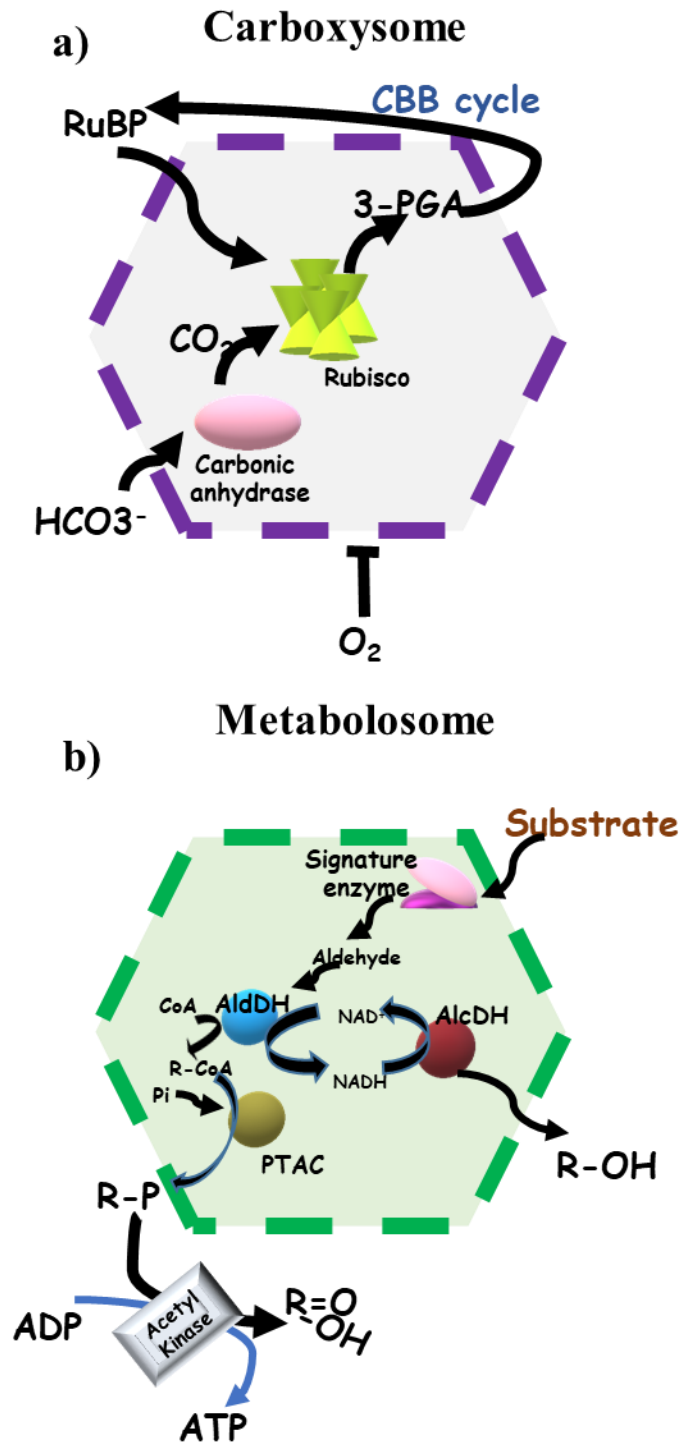
Fig. 1.2: Schematic representation of (a) carboxysome and (b) metabolosome.

1.4.2 Metabolosomes

Many bacterial species contain metabolosomes [21],[22], which are tiny proteinaceous compartments involved in catabolic reactions (**Fig. 1.2b**). Primary job of a metabolosome is to metabolize a specific organic substrate in order to help bacteria thrive in environments

with limited energy sources. Contrary to carboxysomes, metabolosomes are known to metabolize a wide range of organic compounds, and the function of a metabolosome is based on the substrate it metabolizes. For example, propanediol utilization microcompartment catabolizes 1,2-propanediol as a substrate whereas ethanolamine utilization microcompartment catabolizes ethanolamine as a substrate. Every metabolosome has am encapsulated signature enzyme which catalyzes the first step of the cascade reactions. A typical metabolosome is made of both shell proteins and enzymes. The shell proteins allow transport of substrate into the lumen of the compartment where it is converted by the signature enzyme into an aldehyde intermediate.

The aldehyde is then the aldehyde intermediate is then further converted into an alcohol-by-alcohol dehydrogenase that comes out of the compartment. This reaction is paralleled with another reaction where aldehyde dehydrogenase and phosphotransacylase converts the aldehyde dehydrogenase into a phosphorylated product. This requires coenzyme A and recycling of NAD^+ into NADH for alcohol dehydrogenase enzyme. The phosphorylated product comes out of the microcompartment and is dephosphorylated to generate ATP.

Table 1.1: Comparison between α and ß-carboxysome		
	α – carboxysome	*ß – carboxysome*
Gene organization	Core genes located in *cso* operon.[23, 24]	Core genes located in *ccm* cluster as well as in other satellite loci.[23, 24]
Shape	Icosahedral (20 triangular faces)[25]	Icosahedral (20 triangular faces)[26]
Size	Smaller (88-108 nm)[25]	Bigger than α – carboxysome 114-137 nm[26] , 200-400 nm[24, 27]
Outer shell thickness	3-4 nm[25] ≈4 nm[28]	Thicker than α – carboxysome 5-6 nm[29]
Outer shell layer	The outer shell is composed of CcoS1, CsoS2, CsoS3, CsoS4 proteins.[30]	The outer shell layer is composed of CcmK, CcmL and CcmL, CcmO shell proteins.[30]
Inner shell layer	No evidence of inner shell layer. CcmM and CcmN homologues not found.[24]	Inner shell layer is composed of CcmM and CcmN proteins[30, 31]
Facet shell protein	CcoS1 (A-E) hexamer[32]	CcmK hexamer[27, 32]
Vertex shell protein	CsoS4 (A-B) pentamer[32]	CcmL pentamer[27, 32]
Inner lumen density	Loosely packed (low density) inner lumen[30, 33]	Highly dense and highly Para crystalline.[31]

RubisCo arrangement	RuBisCo molecules are arranged in concentric layers. Distance between the first RuBisCo layer and outer shell layer is 6 nm.[25]	RuBisCo molecules are arranged in concentric layers. Distance between the RuBisCo layer and outer shell layer is 2 nm.[31]

1.5 Sixty years of microcompartment research: a journey through the lens of biophysics

It all began during early 1950s, when researchers noted the presence of polyhedral inclusion bodies in the cytoplasm of cyanobacteria, when viewed under electron microscope[34]. These polyhedral bodies, initially thought to be viruses, were later confirmed to be tiny compartments filled with functional enzymes. Using electron micrograph images, Shivley et al., showed the presence of enzyme RubisCo inside these polyhedral compartments found in *Thiobacillus neapolitanus*[35]. These polyhedral structures were called carboxysomes and their identification marked the discovery of the first bacterial microcompartments (MCP). They were small in size (100-200 nm in diameter), made up exclusively of proteins and participated in CO_2 fixation[35]. It took almost 40 years to discover homologues of carboxysomes in other species of bacteria. Functionally diverse MCPs such as 1,2-propanediol microcompartment (PduMCP) and Ethanolamine utilization microcompartments (EutMCP), were discovered in Salmonella sp. [36, 37] [37] and later in other species [38, 39] [39]. Subsequent studies led to the discovery of glycyl-radical dependent MCP [40][41] and MCP involved in ethanol/acetate metabolism[42] and choline fermentation[43]. These MCPs that aid in optimization of metabolic pathways have now been categorized as metabolosomes.

Unlike carboxysomes, which perform CO_2 fixation and aid in photosynthesis, metabolisms provide selective advantage to the bacteria by enabling them to utilize carbon compounds such propanediol/ethanolamine as energy source. They also help in the sequestration of toxic volatile metabolites, protecting the bacteria against cytotoxicity[44].The constituent proteins of MCPs are encoded by genes belonging to a single operon. The genes in the operon encode for shell proteins and enzymes. The shell proteins self-assemble and form the outer coat of the MCP and enzymes are confined within (see reviews by Chowdhury et al 2014 [45]and Kerfeld et al 2018 [13]). Gene deletion/mutation studies have proven to be useful in deciphering the role of individual proteins in MCP. For example, deletion of shell proteins PduA/J alters the size and shape of PduMCP, while deletion of shell protein PduBB' disrupts PduMCP formation[46]. A point mutation in the N-terminal of shell protein PduBB' impairs the encapsulation of enzymes within PduMCP[47]. Similar deletion studies on carboxysomes have shown that proteins CcmM and CcmN are required for the formation of functional carboxysomes[48],[49]. While these studies have helped us to identify proteins (or regions within a protein) essential for the

formation of MCP, they do not yield information regarding the hierarchy in which the MCP proteins assemble, and factors that control their self-assembly. Information regarding the self-assembly of MCP proteins is required not only for understanding the biogenesis of MCPs but also for engineering MCP inspired nano-reactors. The advent of biophysical techniques has enabled us to study the properties and organization of MCPs in detail. A schematic of timeline of research in the field of microcompartment has been shown in **Fig. 1.3.**
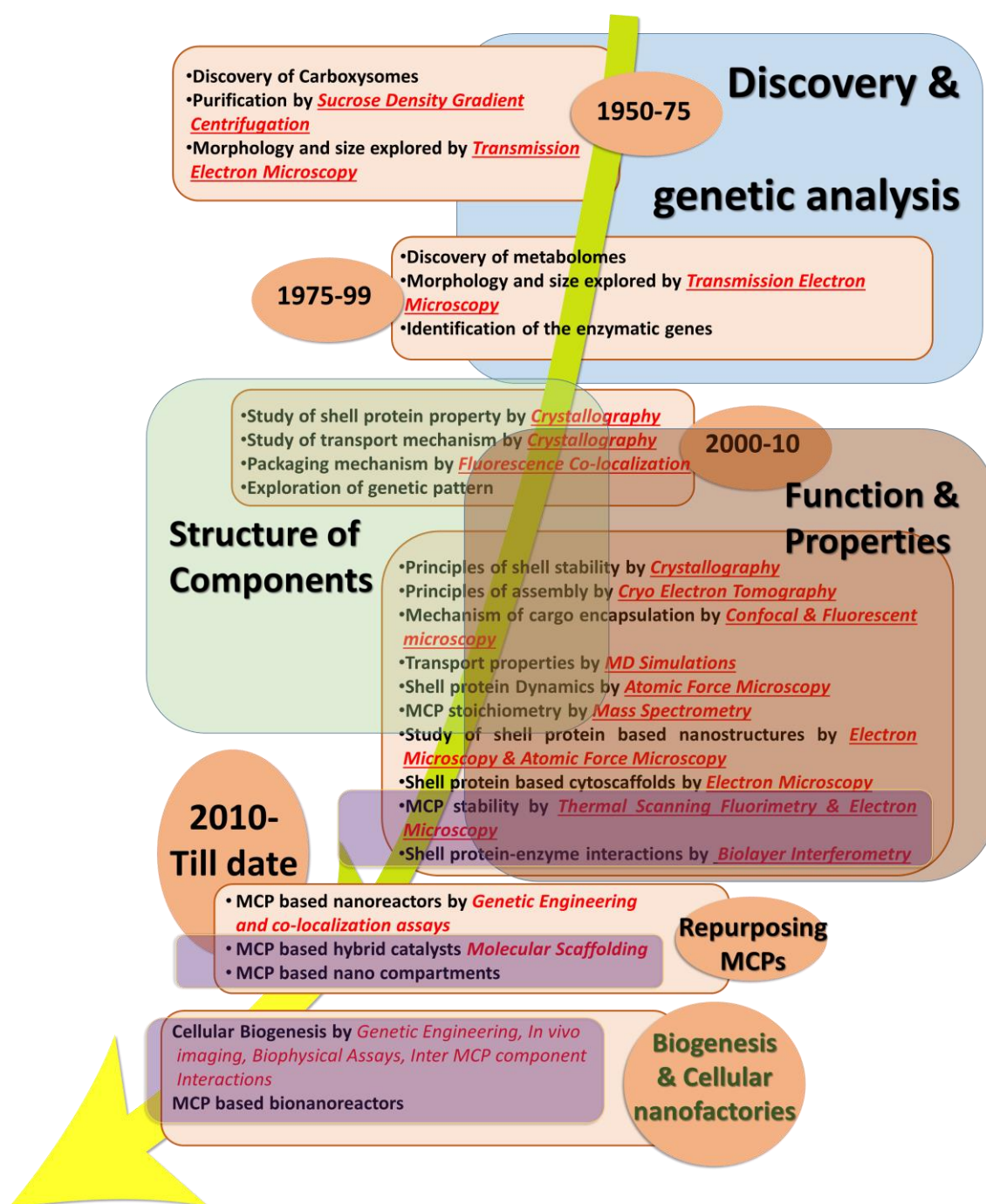
Fig. 1.3: Time line of bacterial microcompartment research (adapted from Kumar and Sinha 2021)[50]

1.5.1 Size, shape and organization

Bacterial microcompartments (MCP) are small and compact assemblages of proteins. Although they resemble viral capsids, they do not show any structural homology with viruses. The extensive use of electron microscopy (EM) has contributed to our understanding regarding the size and morphology of MCPs. Carboxysomes, the first MCPs to be discovered in cyanobacteria are icosahedral in shape and vary in size between 100-150 nm. CryoEM of α-carboxysomes from *Halothiobacillus neapolitanus* showed 88-106 nm sized compartments with icosahedral symmetry of the shell[25]. Similarly, electron microscopy of ß-carboxysome from Synechococcus sp. showed icosahedral compartments of 114-137 nm in size[26]. In a recent report, high resolution TEM has revealed the shell architecture and internal organization of ß-carboxysome purified from Synechococcus sp[31]. These carboxysomes have an average size of 149.9 nm with an edge length (vertex to vertex) of around 72 nm. They have outer shell of 4.5 nm thickness. Internal lumen is highly Para crystalline with a periodic arrangement of Rubisco molecules. This is in contrast with the lumen of α-carboxysomes that are loosely dense. A gap of 2 nm exists between the outer shell and internal Rubisco arrays. This gap is suggested to be occupied by proteins that link outer shell to the Rubisco in the lumen.

The apparent size of the MCPs observed may also depend on the imaging techniques used in study. This is evident from a very recent work published by Kennedy et al.[51]. In this work, the researchers have used imaging techniques like TEM, SEM and CryoEM to observe intact PduMCP purified from *Salmonella sp*. When TEM was used, the PduMCPs appeared to be deflated and were 102 ± 17 nm in diameter. When the samples were subjected to critical point drying and observed under SEM, the PduMCPs appeared to be inflated and with diameter around 126 ± 17 nm. CryoEM captured largest size PduMCPs with diameter around 138 ± 21 nm. However, due to lack of any contrast agent no information could be obtained regarding the surface of the visualized PduMCPs. It must be noted that the intact MCPs along with shell proteins and enzymes are highly complex and big in size (600 MDa). Simplifying the MCP structure by disassociating shell proteins from its enzyme core is very effective while applying EM techniques to MCPs. In 2017, Sutter et al. solved the structure of an intact 6.5 MDa shell prepared heterogeneously using only the shell proteins (Hexamer BMC-H, Trimer BMC-T and Pentamer BMC-P) from Haliangium ochraceum [52]. The 40 nm intact shell composed of 12 pentamers, 60 hexamers and 20 trimers. The shell proteins BMC-H, BMC-P and BMC-T1 were arranged in a single layer. The trimers BMC-T2 and BMC-T3 formed

stacked dimers of trimers. The outer layer of the intact shell had a thickness of 20-30 nm with stacked trimers (BMC-T2 and BMC-T3) protruding outside. The pentamer BMC-P had a pyramid like shape with its base facing outside. The concave side of BMC-H, BMCT1 and lower trimer of stacked BMC-T2 or BMC-T3 faced outside. Overall, the crystal structure of the intact shell gave an insight into the arrangement and orientation of shell proteins on the surface of MCP.

1.5.2 Composition of bacterial microcompartments

Quantification of the protein content of the MCPs is necessary for understanding their stoichiometric composition and identifying the optimum levels of proteins required for designing MCP based synthetic catalytic reactors. This would require a rapid and high throughput techniques. Kim et al 2014 suggested the use of flow cytometry for relative quantification of proteins encapsulated into MCPs[53]. Fusion of signal peptide of PduP enzyme with GFP protein, allowed the encapsulation of the fluorescent protein within PduMCP in Salmonella LT2. Flow cytometry of the bacterial cells containing fusion proteins showed a shift in the fluorescence of the cells grown in the presence of 1,2-PD. No such shift was observed when the cells were grown in the absence of 1,2-PD. By calculating the ratio of the fluorescence of the cells grown in the presence and absence of 1,2-PD, they quantified the level of protein encapsulated within PduMCP. This method also helped them to identify the mutations in the signal peptide that could enhance encapsulation. For example, mutations such as T3A, S4A, R11A, and T12A in the N-terminal of PduP resulted in higher fluorescence of the cells with T3A mutant displaying the highest fluorescence. Overall, flow cytometry method proved to be better than other methods such as fluorescence microscopy, as it enabled quantification of proteins rather than giving a qualitative data.

Recently Yang et al reported the stoichiometric composition of the PduMCP proteins using Qcon-cat driven quantitative mass spectrometry. For this the authors expressed isotopically labelled QconCat peptides (a concatenation of nominated peptides of from all the Pdu proteins) in a cell free system [54]. The purified QconCat peptides were mixed with isolated PduMCPs, and subjected to trypsin digestion. The LC-MS analysis of the digested products allowed absolute quantification of the Pdu proteins. The authors found that the hexamer PduJ to be the most abundant protein in PduMCP and its level increased upon deletion of PduA. Deletion of PduA also resulted in reduction in the level of other shell protein such as PudBB' and enzyme PduCDE. The level of PduP was not affected by PduA deletion. The

quantification of the protein content of PduMCP helped the authors to propose an organizational model for PduMCP. In this model the hexamers PduA and PduJ form the facets of MCP and trimer PduBB' forms the edges.

1.5.3 Assembly of shell proteins

The outer cover of the microcompartments are formed by sheet forming shell proteins. The edge to edge interactions between the shell proteins are crucial to the assembly of the outer shell [55] [56]. Hence, understanding the self-assembly properties of shell proteins and determining their governing factors is a major requisite for studying the self-assembly of microcompartments. In this context, the contributions made by Imaging techniques such as TEM and high-resolution AFM are very significant. The shell proteins have been observed to form tubular structures under both *in vivo* and *in vitro* conditions. Overexpression of hexamer shell protein PduA in E. coli results in the formation of nano-tubes which could be visualized by TEM analysis of the longitudinal section of *E. coli* cells[56]. Both hexamer PduA and trimer PduB have the ability to form nano tubes in vitro when purified at high concentrations (≈5 mg/ml). In the tubular assembly, the shell proteins tend to have their concave side facing outside[57]. Another hexameric homologue, RmmH shell protein from *Mycobacterium smegmatis* self- assembles in to nano-tubes upon overexpression in *E. coli* [58]. TEM image of this shell protein suggests that the concentration of protein, ionic strength and pH of solution determine the higher-order assembly status of shell proteins. At 4 mg/ml RmmH form nano-tubes that tend to disassemble upon dilution to 0.4 mg/ml. Increasing the ionic strength to 500 mM and pH to 10 also disassembles the nano-tubes. The model presented in this study also suggests a concave-out orientation of shell proteins in tubular assembly.

While TEM helps us to visualize the higher-order assemblies of shell protein, it gives no information regarding the dynamics of shell protein assembly due to the drying of the protein samples. This problem has been overcome by visualizing the shell proteins in solution phase using HS-AFM. Sutter et al 2016 used a combination of HS-AFM and X-ray crystallography to visualize the dynamics of shell assembly using hexamer BMC-H from *Haliangium ocraceum[59]*. The crystal structure of this shell protein solved up to 1.8Å resolution show both concave-concave and convex-convex stacking. When observed under HS-AFM, BMC-H formed extended sheets. The 2D patches of sheets had a thickness of ≈ 3.5 nm indicating the formation of the patches by a single layer of shell protein BMC-H. Higher magnification revealed the straight edges and regular angles of the hexamers. Time-lapse HS-AFM at 17

s/frame showed the association and dissociation of hexameric units in to the patches confirming the dynamic nature of shell protein assembly. The mutant variants of BMC-H (K28A) showed larger patches while R78A showed smaller patches. K28A variant also had the tendency to form double layers mediated by convex-convex interactions. Faulkner et al 2019 studied the effect of ionic strength and pH on the dynamic of BMC-H shell proteins using HS-AFM. They observed that the patch size increased with the increase in pH from 3 to 10, suggesting that higher pH favors self-assembly of BMC-H. The dynamic rate of the patches was highest at pH 7, suggesting flexible sheet assembly at neutral pH. An increase in ionic strength also increased the patch size and multilayered sheets could be seen at 500 mM ionic strength[60]. This suggests that salts, particularly divalent cations (Ca^{2+}, Mg^{2+}) are chemical stimulants with the potential to trigger the self-assembly of BMC-H shell proteins.

1.5.4 Physical properties of bacterial microcompartments

Deciphering the complex assembly of MCP would require understanding of the properties of its components. Knowledge about the physical properties of BMC in terms of stability or rigidity and forces that hold the assembly together is essential for the development of BMC inspired nano-bio reactors. Bari et al 2020 performed spectroscopic studies on the structure and stability of PduMCP by using SYPRO Orange dye as a fluorescent probe. SYPRO Orange is a hydrophobic molecule that binds to the hydrophobic patches of proteins when they unfold upon thermal denaturation. The binding of SYPRO Orange to the hydrophobic patches results in increased fluorescence and gives us an idea about the melting temperature of proteins. This study revealed that the core of PduBMC is hydrophobic in nature and hydrophobic interactions between shell protein and enzyme might play a significant role in the PduMCP assembly. The thermal stability of shell proteins is higher than that of encapsulated enzymes and the encapsulation within PduMCP provides stability to the enzymes under thermal stress. Both PduMCP and shell proteins showed an unfolding around 60⁰C, suggesting destabilization of the outer shell around this temperature[61]. This supports an earlier report by Kim et al 2014 where the authors studied the effect of temperature on the structural integrity of PduMCP. TEM images of PduMCP subjected to different temperatures suggested that PduMCPs are stable up to 60⁰C [51].

Mayer et al 2016 performed nanomechanical mapping using atomic force microscopy and compared the stiffness of wild type PduBMC (from *C. freundii*) with that of empty PduBMC (BMC generated using only shell proteins)[62]. Both the wild type and empty BMCs could

tolerate force between 200 pN and 2 nN. Increasing force beyond 2 nN resulted in deformation of both wild type and empty BMC. However, the deformation was higher in case of wild type BMC. This suggested that complete BMCs have lower stiffness than empty MCP. AFM based nano-indentation has also been performed on β-carboxysome from *Synechococcus elongatus* PCC7942 (Syn7942) [31]. A comparison was made between the stiffness of β-carboxysomes and viruses as carboxysomes resemble the icosahedral viral capsids. Unlike viruses, application of force above 300 pN did not cause any break or rupture in β-carboxysome. The spring constant calculated for the β-carboxysomes were found to be ~20 pN nm^{-1}, which is less than that of viruses (40–1250 pN nm^{-1}). This suggests that β-carboxysomes are much softer and less rigid than viral capsids.

1.5.5 Biogenesis of bacterial microcompartments

Over recent years, extensive use of fluorescence microscopy in combination with molecular biology have contributed towards our understanding of the formation and genesis of MCPs. Fluorescence microscopy studies on carboxysomes have not only helped us to visualize the localization of carboxysome proteins in bacterial cell but also revealed the order in which the proteins self-assemble. By fusing CcmN protein and Rbcl (large subunit of RuBisCO) with yellow fluorescent protein and cyan fluorescent protein respectively, Kinney et al 2012 demonstrated that the two proteins co-localize in to fluorescent puncta during carboxysome formation, indicating that CcmN was a part of β-carboxysome [49]. The authors also showed using TEM imaging that CcmN deleted strain could not produce β-carboxysome. Similarly, Cai et al 2013 confirmed that CcmP protein was a part of β-carboxysome. CcmP fused with cerulean fluorescent protein and Rbcl fused with green fluorescent protein co-localized in to fluorescent puncta during β-carboxysome formation [63]. Time lapse fluorescence microscopy revealed that the core of β-carboxysome, comprising RuBisCO, CcmM and CccN proteins, is formed first, to which the shell proteins (CcmK/CcmO) are added[64].

Recently, Wang et al 2019 have shown that RuBisCO interacts CcmM and undergoes condensate formation. Mixing of the two proteins results in the formation of protein droplets. Fluorescence recovery after photobleaching (FRAP) of the protein droplets confirmed their liquid nature [65]. In another study on α-carboxysome, RuBisCO was shown to interact with the N-terminal domain (NTD) of Cso2 protein and undergo phase separation[66]. The affinity of RuBisCO towards the NTD of Cso2 was confirmed using biolayer interferometry, a label

free optical technique to study protein-protein interaction. Mixing RuBisCO with NTD-GFP fused protein resulted in droplet formation as observed under fluorescence microscope [66]. These results suggest that liquid-liquid phase separation could be a driving force behind the assembly of carboxysomes.

1.6 Research paradigm: 1,2-Propanediol utilization microcompartment

Under energy-deficient conditions, certain bacterial species produce 1,2-propanediol utilization microcompartments (PduMCP), which enable bacteria to use 1,2-propanol as their only carbon source. PduMCPs have major implications on the pathogenesis of bacteria including *Salmonella enterica*, when they colonize the lower gut environment. When plant sugars like rhamnose and fucose are broken down in the gut, 1,2-propanediol is produced as a byproduct.1,2-propanediol induces the production of PduMCP proteins (shell proteins and enzymes) from genes encoded by *pdu operon*. The proteins then self-assemble to create compartment-like structures, with shell proteins serving as the exterior envelope and enzymes contained inside. **Fig. 1.4** gives a schematic representation of PduMCP from *Salmonella enterica* LT2.

The outer shell of PduMCP is made of 8 shell proteins: PduA, B/B', J, M, N, K, T and U. The shell proteins PduA, PduB, PduB', PduJ, PduK, PduT, and PduU are hexagonal in form and are members of the family of bacterial microcompartment (BMC) domains. PduN is a pentagonal BMV protein that makes up the vertices of MCPs. Among the BMC domain proteins, PduA, PduJ and PduU are hexamers and six BMC-H domains associate to form one hexagon. PduB/PduB' and PduT are trimers where two BMC-H domains are fused in tandem to form BMC-T domain. Three BMC-T domains come together to form one hexagon. PduN on the other hand is made of five BMC-P domain that self-assemble to form one pentagon. The core of PduMCP is composed of enzymes involved in the metabolism of 1,2-propanediol. The substrate 1,2-propanediol enters the compartment (likely through the pore of PduA) and is converted into propionaldehyde by the signature enzyme diol dehydratase PduCDE. The toxic/volatile intermediate propionaldehyde is retained inside and is further metabolized to propanol by alcohol dehydrogenase PduQ in a reaction that regenerates NAD^+ from NADH. The NAD^+ is utilized by aldehyde dehydrogenase PduP which converts propionaldehyde to propionyl-coenzyme A. It is further phosphorylated by
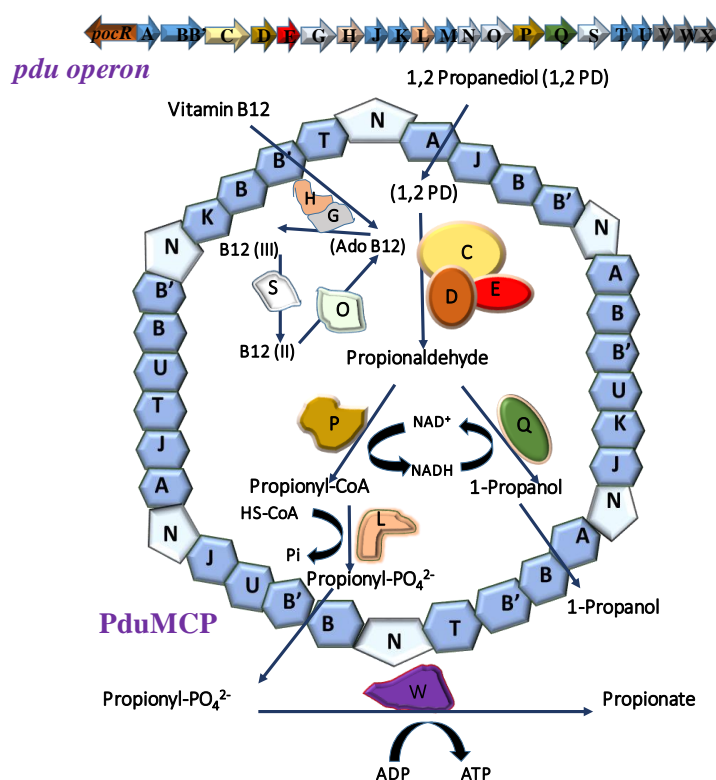
Fig. 1.4: Schematic representation of *pdu operon* and PduMCP (adapted from Kumar et al., 2022) [67]

phosphotransacylase PduL into propionyl-phosphate, which comes out of PduMCP where it is dephosphorylated to generate ATP. The energy generated during 1,2-propanediol metabolism is helps *Salmonella enterica* LT2 to sustain where there is a scarcity of food in the surrounding environment.

1.7 Research questions

Since the microcompartments have endured throughout evolution, it is indisputable that they have offered bacteria a survival benefit. Therefore, it's crucial to comprehend how enclosing enzymes in a compartment might benefit the metabolic processes. We have understood from the reports that are currently available in the literature that compartmentalization protects bacteria from the toxicity of volatile intermediate products. In the case of PduMCP, for instance, the pore of PduA permits the entry of 1,2-propanediol for its metabolism but inhibits the removal of propionaldehyde, a hazardous volatile intermediate [44]. In this manner, the bacterial outer shell prevents cytotoxicity and promotes bacterial survival. It is unclear, nevertheless, if the benefit of the outer shell is limited to defense against hazardous intermediates. For example, shell protein PduA has been demonstrated to transport 1,2-

propanediol to the lumen of PduMCP while preventing the hazardous intermediary propionaldehyde from escaping. We therefore ask the following questions:

- What benefits other shell proteins provide the bacteria?
- Why have bacteria evolved to accommodate many types of shell proteins in the outer shell if the restriction of harmful metabolites is the exclusive function of the outer shell? Do other shell proteins have different purposes?
- Does compartmentalization endow additional advantages to the metabolic process within the microcompartment?

In chapter 3 of his thesis, I have addressed these questions using a barrage of biochemical and biophysical approaches. I have used the major shell protein and the signature enzyme of PduMCP to investigate how the shell protein might influence the functioning of the native enzyme and what implications this would have on the intact PduMCP.

Although it is crucial to study how enzyme compartmentalization and shell proteins affect processes, one cannot help but wonder how the individual proteins that make up a microcompartment come together in the first place. The mechanism by which the microcompartments are formed could be understood by figuring out what causes these proteins to bind together. We have a fairly good understanding about the ability of shell proteins to form higher order assemblies. In an aqueous medium, they can form extended sheets by participating in edge-to-edge interactions. However, the proteins must nucleate and come close to one another in order to form compartment-like structures. Besides, the self-assembly behavior of enzymes in association with shell proteins is not well understood. This necessitates the question of what underlying factors cause the shell protein and enzyme to connect and self-assemble. This raises the following questions:

- How would change in the environmental conditions affect the self-assembly of shell protein and enzyme?
- How would the self-assembly of shell protein and enzyme impact the catalytic proficiency of the enzyme?

I have addressed the queries in Chapter 4 of his thesis. By alerting the conditions of the surrounding environment of shell protein and enzyme, I have elucidated the factors that govern the self-assembly mechanism of shell protein and enzyme in solution phase.

The formation of a microcompartment would involve an association between shell protein and enzyme. Beside understanding the affinity between shell protein and enzyme, it is also important that we regions involved in the shell protein and enzyme interaction is elucidated. An interesting characteristic of membraneless organelles is that they are rich in disordered regions. These disordered regions enable allow multivalent association between the partner proteins and promote self-assembly. While microcompartments are devoid of lipid bilayer, they do have an outer layer of shell protein that form ordered facets. Determining if disordered regions in microcompartment proteins exist would be a crucial question to answer. Notably, it has been shown that disordered regions are crucial for the formation and self-assembly of carboxysomes [66]. This information motivates us to ask the following questions:

- Do metabolosomes also contain disordered regions?
- What role do disordered regions play in the interaction between shells and enzymes?

In Chapter 5 of his thesis, I have used a combination of bioinformatics, computational, and biochemical methods to study the distribution of disordered regions in PduMCP and their probable role in shell enzyme interaction.

Note:

The permission has been granted by authors and corresponding author of the published paper prior to its adoption in the present thesis. The publication associated with this work is:

Kumar G, Sinha S. Biophysical approaches to understand and re-purpose bacterial microcompartments. *Current Opinion in Microbiology*. 2021 Oct 1; 63:43-51.

Chapter 2

Materials and Methods

Biochemistry

Molecular biology

Bioinformatics computational

Biophysics

Imaging
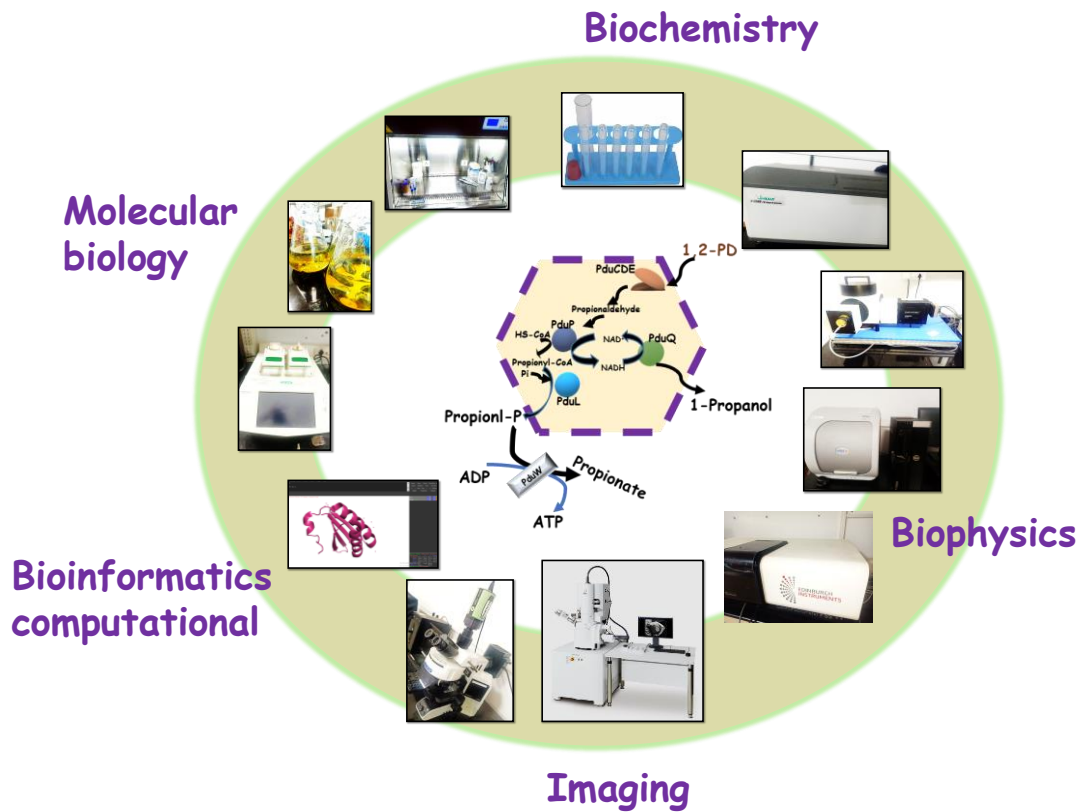
In my thesis, I have explored UV-Visible spectroscopy to perform enzyme assays and kinetics. The thermal stability studies of microcompartment, shell protein and enzyme has been performed using fluorescence spectroscopy, circular dichroism and differential scanning fluorimetry. These techniques helped me to figure out the significance enzyme encapsulation within microcompartment. Techniques such as dynamic light scattering and turbidity assays were helpful in probing the self-assembly and identifying the factors governing the self-assembly of shell proteins in solution phase. To visualize the dynamics of shell protein and enzyme assembly in solution, I used florescence microcopy. I learned how to label proteins with commercial fluorophores and study self-assembly behavior of shell proteins and enzyme. The morphology of the purified microcompartment, shell protein and enzyme has been studied using electron microcopy. I used the overlap PCR technique to generate point mutations and truncated variants of shell protein and enzyme followed by cloning genes into expression vector and expression of proteins in *E.coli* BL21(DE3) cells. I've always felt it was helpful to employ computational modelling and bioinformatics techniques to complement my experimental findings. They gave me insight into the residual level of interactions between shell protein and enzyme and helped me better grasp the functions of the proteins' structured and unstructured regions. In this chapter I detail out the materials used and the methodology I have followed for my study.

2.1 Materials

Table 2.1: List of chemicals and reagents used in the present thesis	
Name of the chemical/reagent	**Company**
3-methyl-2-benzothiazoline hydrazine	Sigma Aldrich, India
Acrylamide	Hi-Media, India
Ado-Cobalamin	Sigma Aldrich, India
Ammonium persulphate	Sigma Aldrich, India
8-Anilinonaphthalene-1-sulfonic acid	Sigma Aldrich, India
B-PER II, Bacterial protein extraction reagent	ThermoFischer Scientific, India
Bis-acrylamide	Sigma Aldrich, India
Bradford reagent	Sigma Aldrich, India
Calcium Chloride	Sigma-Aldrich, India
Citric acid monohydrate	Himedia, India

Cyanocobalamin	Sigma Aldrich, India
DMSO	Sigma Aldrich, India
DNase	Sigma Aldrich, India
EDTA disodium salt dihydrate	Hi-Media, India
Ethanol	Hayman, UK
Glycine	Hi-Media, India
Imidazole	Hi-Media, India
IPTG	Hi-Media, India
Luria Broth	Hi-Media, India
Lysozyme	Hi-Media, India
Magnesium chloride	Hi-Media, India
Magnesium sulphate heptahydrate	Hi-Media, India
NHS ester-Alexa 488	Tocris Bioscience, UK
NHS ester-Texas red	Tocris Bioscience, UK
Ni-NTA resin	Qiagen, Thermo Fischer Scientific, India
Nickel sulphate hexahydrate	Sigma Aldrich, India
PEG-6000	Sigma Aldrich, India
PMSF	Sigma Aldrich, India
Potassium Chloride	Hi-Media, India
Potassium Hydroxide	Hi-Media, India
Potassium phosphate monobasic anhydrous	Hi-Media, India
Propane-1,2-diol	Sigma Aldrich, India
Sodium Azide	Merck, India
Sodium Chloride H	i-Media, India
Sodium Do-decyl sulphate	Hi-Media, India
Sodium Hydroxide pellets	Hi-Media, India
Sodium phosphate dibasic anhydrous	Hi-Media, India
Sodium phosphate monobasic anhydrous	Hi-Media, India
Succinic acid disodium salt	Sigma Aldrich, India
TEMED	Sigma Aldrich, India
Tris Base	Hi-Media, India

Table 2.2: List of molecular biology materials and enzymes	
DNTPs	NEB, UK
Vent polymerase	NEB, UK
Taq polymerase	NEB, UK
Calf intestinal phosphatase	NEB, UK
Restriction enzyme NdeI	NEB, UK
Restriction enzyme XhoI	NEB, UK
Restriction enzyme BglII	NEB, UK
Restriction enzyme HindIII	NEB,UK
Agarose	Lonza, Switzerland
ETBR	Sigma Aldrich, India
T4 DNA ligase	NEB, UK

Table 2.3 List of the primers used for generating constructs		
PduBM38L	Forward primer 1(Nde1 sequence):	5'-gaaccgcatatgcatcaccatcaccatcat agcagcaatgagctg-3'
	Reverse primer 1 (Xho1 sequence):	5'-aggcggctcgagtcagatgtaggacggacgatc-3'
	Forward primer 2:	5'-cgagagacggctctggcagaaaaaagc-3'
	Reverse primer 2:	5'-gctttttctgccagagccgtctctcg-3'
PduC(ΔD)E	Forward primer 1 (Bgl II sequence):	5'-gccaatagatctatgcatcaccaccatcatcacatgag atcgaaaagatttgaagcactg-3'
	Reverse primer 1 (Hind III sequence):	5'-aaa ctt aag ctt gaa atc ctt aat cgt cgc ctt tga gtt ttt tac gct caa cgt aca gcc-3'
	Forward primer 2:	5'-attaaggggtgagaaatgggcagcgataaacccgtctc-3'
	Reverse primer 2:	5'-gagacgggtttatcgctgcccatttctcacccccttaat-3'

Table 2.4 List of bacterial strains used in the present thesis	
Salmonella enterica serovar Typhimurium LT2	Gift from Dr. TA Bobik, Iowa State University
E. coli BL21(DE3)/ pET41a –PduBB'	Gift from Dr. TA Bobik, Iowa State University

E. coli BL21(DE3)/ pET41a –PduB'	Gift from Dr. TA Bobik, Iowa State University
E. coli BL21(DE3)/ pET41a –PduA	Gift from Dr. TA Bobik, Iowa State University
E. coli BL21(DE3)/ pET41a –PduBM38L	Generated in Sinha Lab
E. coli BL21(DE3)/ pET41a –PduCDE	Generated in Sinha Lab
E. coli BL21(DE3)/ pET41a –PduC(ΔD)E	Generated in Sinha Lab

2.2 Methodology

2.2.1 Purification of 1,2-propanediol utilization microcompartment (PduMCP)

1% of overnight grown culture of *Salmonella enterica* LT2 was inoculated in 400 ml of 1X NCE (non-carbon E) medium, which was supplemented with 0.6 % 1,2-Propanediol, 0.5 % succinic acid and 1mM of $MgSO_4$ and incubated at 37°C for 16 h. The harvested cells were washed with buffer A (containing 50 mM Tris Base pH 8, 500 mM KCl, 25 mM NaCl, 12.5 mM $MgCl_2$, 1.5 % 1,2-Propanediol) at 8000 X g for 5 min at 4 °C. The washed cells were then re-suspended in Buffer A containing 75% bacterial protein extraction reagent (BPER-II), 2 mg DNase, 0.4 mM phenylmethane sulfonyl fluoride (PMSF) and 1 mg/ml of lysozyme. The re-suspended cells were kept on a shaker at 45-50 rpm at room temperature for 30 min, followed by incubation on ice for 5 min. The lysed cells were removed by centrifugation at 12,000 X g for 5 min at 4 °C and PduMCPs in the supernatant are further pelleted by centrifugation at 20000 X g for 20 min at 4 °C. The pellet was re-suspended in Buffer A containing 60% of B-PER II and 0.4 mM PMSF and was centrifuged at 20,000 X g for 20 min at 4 °C. The supernatant was discarded and the thin film obtained was re-suspended in pre-chilled Buffer B (containing 50 mM Tris Base pH 8, 50 mM KCl, 5 mM $MgCl_2$, 1 % 1,2-Propanediol). The re-suspended thin film was centrifuged at 12,000 X g for 5 min at 4 °C . The supernatant containing PduMCP was collected and stored at 4°C for further use. The concentration of PduMCP was estimated using Bradford reagent and purity was tested by performing SDS-PAGE.

2.2.2 Expression and purification of PduCDE and shell proteins PduBB' and PduB'

1% of overnight grown primary culture of *E.coli* BL21(DE3) (transformed with genes for PduCDE or shell proteins cloned into pTA925 vector, containing Kanamycin resistant gene) was inoculated in 400 ml of LB media (containing 50 μg/ml kanamycin) and incubated at

37°C for 1.5 to 2 h until it reached an OD_{600} of 0.5. PduCDE expression was induced by adding 1 mM IPTG and incubating the culture at 37°C for 4 h. PduBB'/PduB' expression was induced by adding 0.5 mM IPTG and incubating the culture at 28 °C for 12 h. The cells were harvested and lysed in column buffer (50 mM Tris-base pH 7.5, 200 mM NaCl and 5 mM imidazole). The supernatant was passed through Ni-NTA column. The column was then washed with washing buffer (50 mM Tris-base pH 7.5, 200 mM NaCl, 50 mM imidazole) and protein of interest was eluted by passing elution buffer (50 mM Tris-base pH 7.5, 200 mM NaCl, 200 mM imidazole). The eluted protein samples were dialyzed using 10 kDa cutoff dialysis membrane in 10 mM sodium phosphate buffer (pH 7.4). Protein concentration was estimated using Bradford reagent and purity was checked by performing SDS-PAGE.

2.2.3 Cloning and expression of PduBM38L

To get full length PduB (270 amino acids), we generated a single mutant PduBM38L by creating point mutation ΔM38L using overlap extension PCR with the following primers:

External primers:

Forward primer 1 (Nde1 sequence):

5'-gaaccgcatatgcatcaccatcaccatcat agcagcaatgagctg-3'

Reverse primer 1 (Xho1 sequence):

5'-aggcggctcgagtcagatgtaggacggacgatc-3'

Internal primers for creating point mutation:

Forward primer 2: 5'-cgagagacggct**ctg**gcagaaaaaagc-3'

Reverse primer 2: 5'-gcttttttctgc**cag**agccgtctctcg-3'

In the internal primers we replaced the start codon 'atg' for Methionine with codon 'ctg' (bold and underlined) for Leucin. The point mutation was confirmed by sequencing result. The mutant PCR product with Nde1 and Xho1 restriction sites were cloned into pet21(a) vector having ampicillin resistant gene and transformed into *E. coli* DH5α strain. The positive clones were selected on ampicillin plates and were further inoculated for primary culture. Plasmids were isolated from the primary cultures of DH5α cells and were transformed into *E. coli* BL21(DE3) expression strain.

For protein purification 1% of overnight grown primary culture *E. coli* BL21(DE3) (transformed with PduBM38L) were inoculated in 400 ml of LB media (containing 100 µg/ml ampicillin) and incubated at 37°C for 1.5 to 2 h until it reached an OD_{600} of 0.5. Protein expression was induced by adding 0.5 mM IPTG and incubating the culture at 28 °C for 12 h. The cells were harvested and lysed in column buffer (50 mM Tris-base pH 7.5, 200 mM NaCl and 5 mM imidazole). The supernatant was passed through Ni-NTA column. The column was then washed with washing buffer (50 mM Tris-base pH 7.5, 200 mM NaCl and 50 mM imidazole) and protein of interest was eluted by passing elution buffer (50 mM Tris-base pH 7.5, 200 mM NaCl and 200 mM imidazole). The eluted protein samples were dialyzed in 10 mM sodium phosphate buffer (pH 7.4). Protein concentration was checked using Bradford reagent and purity was checked by performing SDS-PAGE.

2.2.4 Cloning and expression of PduC(Δ2-20D)E

We generated *PduC(Δ2-20D)E* by overlap extension PCR with the following primers:

External primers:

Forward primer 1 (Bgl II sequence):

5'-gcc aat aga tct <u>atg</u> cat cac cac cat cat cac <u>atg</u> aga tcg aaa aga ttt gaa gca ctg-3'

Reverse primer 1 (Hind III sequence):

5'-aaa ctt aag ctt gaa atc ctt aat cgt cgc ctt tga gtt ttt tac gct caa cgt aca gcc-3'

Internal primers for deletion:

5'-attaaggggtgag aaatg <u>gg</u> cagcga taaacccgtc tc-3'

5'-gagacgggtttatcgctgcccatttctcaccccttaat-3'

The mutant PCR product with Nde1 and Xho1 restriction sites were cloned into pTA925 vector having Kanamycin resistant gene and transformed into *E. coli* DH5α strain. The positive clones were selected on kanamycin plates and were further inoculated for primary culture. Plasmids were isolated form the primary cultures of DH5α cells and transformed into *E. coli* BL21(DE3) expression strain.

For protein purification, 1% of overnight grown primary culture *E. coli* BL21(DE3) (transformed with *PduC(Δ2-20DE)* were inoculated in 400 ml of LB media (containing 50 µg/ml Kanamycin) and incubated at 37°C for 1.5 to 2 h until it reached an OD_{600} of 0.5.

Protein expression was induced by adding 1 mM IPTG and incubating the culture at 37 °C for 3 h. The cells were harvested and lysed in column buffer (50 mM Tris-base pH 7.5, 200 mM NaCl and 5 mM imidazole). The supernatant was passed through Ni-NTA column. The column was washed with washing buffer (50 mM Tris-base pH 7.5, 200 mM NaCl and 50 mM imidazole) and protein of interest was eluted by passing elution buffer (50 mM Tris-base pH 7.5, 200 mM NaCl and 200 mM imidazole). The eluted protein samples were dialyzed in 10 mM sodium phosphate buffer (pH 7.4). Protein concentration was checked using Bradford reagent and purity was checked by performing SDS-PAGE.

2.2.5 Diol dehydratase assay

Diol dehydratase activity of purified PduMCP or PduCDE is estimated by 3-methyl-2-benzothiazoline hydrazine (MBTH) method. For routine assays, 2 μg of enzyme (for PduCDE assay) and or 5 μg of PduMCP (for PduMCP assay) is added to 900 μl of the assay buffer (0.2 M 1,2-propanediol, 0.05 M KCl, 0.035 M potassium phosphate buffer (pH 8.0) at 37 °C. Reaction is started by adding 50 μl of adenosyl cobalamin AdoCbl (15 μM) and quenched after 10 min by adding 1 ml of potassium citrate buffer (pH 3.6). Then, 0.5 ml of 0.1 % MBTH (w/v) is added to the reaction mixture and the it is incubated at 37 °C for 15 min. After 15 min, 1 ml of dH_2O is added and absorbance of the product formed is taken at 305 nm using UV spectrophotometer. Absorbance at 304 nm is converted to the concentration of the product formed using molar extinction coefficient 13000 M^{-1} cm^{-1}. We define specific activity as μmol of product formed by 1 mg of PduCDE or PduMCP in 1 min.

2.2.6 Thermal shock assays using PduMCP

To check the effect of temperature on the activity of PduMCP and PduCDE, the proteins samples (50 μl of 0.5 mg/ml of PduCDE or PduMCP, in 0.5 ml micro centrifuge tubes) were exposed to temperatures ranging from 4°C to 60°C for a duration of 2 min. Since, diol dehydratase assays requires a temperature sensitive ado-cobalamin (ado-B12) as a cofactor, it was difficult to perform the enzyme assays at temperatures above 37°C. We therefore opted for thermal shock experiment, where PduCDE or PduMCPs were subjected to thermal shock at specified temperatures followed by equilibrating the samples at 37°C before performing the enzyme assay. For 4°C treatment, we keep the protein samples in 4°C refrigerator for 2 min. For thermal shock at temperatures 25°C to 60°C we placed the protein samples in water bath for 2 min. Then the samples were brought to room temperature and subjected to diol dehydratase assay, as described above in **section 2.2.5**.

To check for the effect of shell proteins on the activity of PduCDE enzyme, 40 µM of shell protein (PduBB'/PduB'/PduBM38L) was mixed with PduCDE at different molar ratio and incubated at 4°C for 1 h (total volume of the mixture was 50 µl). Shell proteins and enzyme were mixed in different molar ratios based on the concentration of shell protein monomers and PduCDE trimeric complex. The samples were then subjected to thermal shock at 37°C , 45°C and 50°C in water bath for a short duration of 2 min. After cooling down the samples to room temperature, they were equilibrated at 37°C followed by MBTH diol dehydratase assay, as described above in **section 2.2.5.**

2.2.7 Fluorophore labeling of shell protein and enzyme

i) Labeling of shell protein with Acrylodan dye

1 mg/ml of shell proteins were labeled with Acrylodan in 5:1 (protein: dye in µM) ratio. The samples were incubated overnight at 4°C. The dye labelled protein was separated from free dye using PD10 columns. The presence of dye was confirmed using UV-Vis spectrophotometer and the presence of protein is confirmed using and Bradford assay. CD spectroscopy was performed to check for the effect of fluorophore labeling on native conformation of protein.

ii) Labeling of shell protein with Alexa-488 and PduCDE with Texas Red

400 µl of shell protein (PduBB'/PduBM38L/PduB') or enzyme PduCDE of concentration 1 mg/ml was mixed with Alexa-488 or Texas Red fluorophores in protein: dye molar ratio of 1:5. The samples were incubated overnight at 4°C followed by excess dye removal by dialysis in 10 mM sodium phosphate buffer (pH 7.4). The dye labeled protein samples were used for microscopic and spectroscopic studies. To get higher concentration of protein, the samples were concentrated using 10 kDa cutoff centrifugal filter units. CD spectroscopy was performed to check for the effect of fluorophore labeling on native conformation of protein.

2.2.8 Fluorescence spectroscopy

i) Temperature dependent fluorescence spectroscopy

Temperature dependent intrinsic fluorescence assay for PduMCP and PduCDE was carried out using Spectrofluorimeter FS5 (Edinburgh Instruments, UK). 400 µl of 0.1 mg/ml of PduMCP or 0.1 mg/ml of PduCDE were excited at 280 nm (band width = 1 nm) and emission spectra (band width = 1 nm) were recorded between 300 to 500 nm. 1st derivative of the

emission maxima (328 nm) versus temperature plot revealed the melting temperatures (T_m values) of PduCDE and PduMCP.

To check the effect of EDTA on the thermal stability of PduMCP, the purified PduMCP samples are incubated in the absence and presence of EDTA at different concentrations for 5 min at room temperature. Temperature dependent intrinsic fluorescence assay for the EDTA treated and control samples of PduMCP is carried out using Spectrofluorimeter FS5 (Edinburgh Instruments, UK). The sample are excited at 280 nm (band width = 1 nm) and emission spectra (band width = 1 nm) are recoded between 300 to 500 nm at temperatures between 20°C-90°C. The 1st derivative of the emission maxima versus temperature plot reveals the melting temperatures (T_m values) of PduMCP.

ii) Fluorescence lifetime measurements

Alexa-488 labeled PduBB' samples are incubated in the presence of different salts (KCl, NaCl, $MgCl_2$, and $MgSO_4$) for 1 h at room temperature. Fluorescence lifetime of the samples is determined using TCSPC system, HORIBA DeltaFlex (Horiba, Japan). The samples are excited using 402 laser diode and emission is recorded at 520 nm at bandwidth of 2 nm. Before recording the lifetime decay of the samples the instrument response factor is determined using Ludox solution. The life time decay curves were best fitted to mono-exponential decay equation to get the lifetime of the samples.

iii) Fluorescence anisotropy measurements for shell protein and enzyme interactions

3 μM of the shell proteins (labeled with Acrylodan) and enzyme [PduCDE or PduC(Δ2-20D)E] were mixed in equimolar ratio and incubated at 4°C for 1 h. Mixing the proteins in equimolar ratio ensured that for every mole of shell protein subunit in solution, one mole of PduCDE complex was available in the surrounding for interaction. Time resolved fluorescence anisotropy measurements were carried out using time correlated single photon counting (TCSPC) (Horiba, Japan) instrument equipped with 375 nm laser (Nd:YAG). 300 μl of protein samples were used in the experiment. We collected the fluorescence emission at 485 nm wavelength at 0° (parallel) and 90° (perpendicular) angle with respect to excitation light. Perpendicular component of the emission was corrected using the G-factor. G-factor was determined independently for the free dye. Instrument response function was found to be ~ 180 ps. Anisotropy decays were fitted using bi-exponential decay model:

$$r(t) = r_0\left(\beta_{fast}e^{-t/\tau_{fast}} + \beta_{slow}e^{-t/\tau_{slow}}\right)$$ -------------------*Eq. 1*[68]

Here r_0 is the intrinsic fluorescence anisotropy, τ_{fast} and τ_{slow} are faster and slower rotational correlation time and β_{fast} and β_{slow} are amplitudes of faster and slower rotational correlation time respectively.

iv) Fluorescence anisotropy measurements to assess conformational flexibility

Alexa-488 labeled PduBB' is mixed with 5 % w/v of PEG-6000 and 1mM $MgSO_4$ to induce phase separation. Time resolved fluorescence anisotropy of the PduBB' samples under control and phase separated condition is determined using TCSPC system, HORIBA DeltaFlex (Horiba, Japan) as described above. Rotational correlation time is determined by fitting the anisotropy decay curve to mono exponential decay equation.

2.2.9 Dynamic light scattering (DLS)

i) Effect of protein concentration on size distribution of shell proteins

The size distribution of the shell proteins (PduBB'/PduB'/PduBM38L) at different concentrations (0.1 – 0.8 mg/ml) were recorded using a ZetaSizer Nano ZSP (Malvern Instruments, UK). The scattered intensity was measured at backscattered angle of 173°. For each sample, three readings were recorded. The size distribution of the three shell proteins was compared by plotting intensity percentage of the samples at different shell protein concentrations.

ii) *Effect of slats on size distribution of shell proteins*

To study the effect of salts on the size distribution of shell protein PduBB', 200 µl of the shell protein samples (0.1 mg/ml) are incubated in the presence of different concentrations of salts (KCl, NaCl, $MgCl_2$, and $MgSO_4$) for 1h at RT. The size distribution of the shell protein is recorded using ZetaSizer Nano ZSP (Malvern Instruments, UK). The scattered intensity at backscattered angle 173° is recorded, and for each sample 3 measurements are recorded. To determine the effect of salts on the surface charge of PduBB', 200 µl of PduBB' solution is incubated in the presence of different salts (KCl, NaCl, $MgCl_2$, and $MgSO_4$) for 1 h at RT. Zeta potential of each sample is measured using ZetaSizer Nano ZSP (Malvern Instruments, UK).

2.2.10 Transmission electron microscopy

TEM imaging of the shell proteins and enzyme was carried out by drop casting 10 µl of 0.2 mg/ml shell proteins (PduBB'/PduB'/PduBM38L) or 10 µl of 0.2 mg/ml PduCDE enzyme on

to carbon coated TEM grids, followed by staining with 7 µl of 1 % (w/v) uranyl nitrate (freshly prepared and filtered). The grids were then washed with Milli-Q and air dried for 24 h at room temperature in dark. TEM images were obtained using JEM 2100 TEM (JEOL, USA) operated at 120 kV.

2.2.11 Circular Dichroism spectroscopy

i) Temperature dependent CD spectroscopy of shell proteins

CD spectra of shell proteins (PduBB'/PduB'/PduBM38L) at different concentrations (0.1 – 0.8 mg/ml) were recorded using a CD spectrophotometer (Jasco J-1500, CD spectrophotometer, (Jasco, Japan). 200 µl of samples were taken in a quartz cell with a path length of (0.1 cm) in a nitrogen atmosphere for measurements in the far-UV region (195-260 nm). An accumulation of three scans with scanning speed of 200 nm/min was performed. Sample temperature was maintained at 25 °C using a mini circulation water bath (Jasco MCB-100) connected to the water-jacketed sample holder chamber. Molar ellipticity was calculated from millidegree by dividing the millidegrees by molar concentration of protein multiplied and pathlength of the cuvette used (1 mm). The path length value and molar concentrations were provided in the Jasco-1000 spectral analysis software and CD spectra with molar ellipticity in the Y-axis was obtained. To check the effect of temperature on secondary structures of shell proteins, CD spectra of 200 µl of shell proteins (0.25 mg/ml) were recorded from 25 °C to 100°C at the rate of 2 °C per min.

ii) Effect of salts on native conformation of shell protein

PduBB' sample incubated in the presence of salts are dialyzed against phosphate buffer (pH 7.4). The dialyzed samples (200 µl) are used for determination of secondary structure of PduBB' protein in a 0.1 cm quartz cuvette in N_2 atmosphere using a CD spectrophotometer (Jasco J-1500, CD spectrophotometer, (Jasco, Japan). For the determination of CD spectra 200 µl of samples are taken in a quartz cell of path length of (0.1 cm) in a N_2 atmosphere. CD spectra are measured in the far-UV region (195-260 nm). Tree scans with scanning speed of 200 nm/min is accumulated for each sample at 25°C.

2.2.12 Biolayer interferometry

Interactions between the enzyme and shell proteins (PduBB'/PduB'/PduBM38L) at different shell protein concentrations (0.1 – 0.8 mg/ml) were carried out using Fortebio Octet K2

(Molecular Devices, USA). Amine reactive second generation (AR2G) biosensors were activated using a solution containing 40mM 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 20 mM of N-hydroxysuccinimide (NHS) mixed in 2:1 molar ratio.   0.5 mg/ml of PduBB' or PduCDE (PduCDE in case of enzyme-shell interaction study) is loaded onto the activated sensor. The enzyme loaded sensors were then immersed in ethanolamine (pH 8.0) to block or mask the sites on sensor tips where enzyme molecules were not present. The sensors were then immersed in titer wells containing shell proteins to record their association kinetics on to the enzyme loaded sensors. The dissociation kinetics were recorded by immersing the sensor tips in 10 mM sodium phosphate buffer (pH 7.4). The sensors were regenerated using 2M NaCl. Global fitting of the association and dissociation kinetics is performed using Data Analysis HT 9.0.0.33 software provided with the OctateK2 instrument.

2.2.13 Aggregation kinetics experiment

To study the aggregation of PduCDE in the presence of shell proteins (PduBB'/PduB'/PduBM38L), 5μM of PduCDE was mixed with 40 μM of shell proteins (only buffer in control sample) and incubated at 4°C for 1 h. To check the effect of PduB'+PduBM38L mixture of PduCDE aggregation, 5μM of PduCDE was mixed with PduB' (20 μM) + PduBM38L (20 μM) mixture and incubated at 4°C for 1 h. To check the effect of a non-native protein on PduCDE aggregation, 5μM of PduCDE was mixed with 80 μM of lysozyme and incubated at 4°C for 1 h. For aggregation studies, protein samples were maintained at 45°C in Spectrofluorometer FS5 (Edinburgh Instruments, UK). The samples are excited at 360 nm, band width of 1 nm and scattered light at 360 nm (band width of 1 nm) are recorded as a function of time.

2.2.14 Thermal shift assay

5μM of PduCDE was mixed with 40 μM of shell proteins (only buffer in control sample; total volume of mixture was 400 μl), and incubated at 4°C for 1 h. 4 μM of 8-Anilinonaphthalene-1-sulfonic acid was added to the 400 μl protein samples and kept at room temperature (25°C) for 5 min. The samples were excited at 377 nm and emissions at 488 nm were recorded at temperatures ranging from 25°C to 80°C.

2.2.15 Molecular docking

De novo modeling of PduD and ΔN-terminal PduD was carried out using Galaxy web server.[69, 70] For flexible docking, PDB file of the structure and amino acid sequence of

PduB were uploaded as input into the CABS dock server.[71] The N-terminal residues of PduD were marked as flexible region and 50 cycles of simulation were run. The residual level interactions in the PduD: N-terminal PduB docked complex were determined using PDB Sum server. The free energy of binding was determined using PRODIGY server.[72]

2.2.16 Turbidity Assay

The turbidity assay is performed with 0.5 mg/ml of shell protein (treated with salts KCl, NaCl, MgCl$_2$, and MgSO$_4$) at room temperature for 1 h) in a 96 well plate using multiplate reader Infinite M Plex (Tecan, Austria), in absorbance mode at 600 nm.

2.2.17 Fluorescence microscopy and visualization of phase separation

For visualization of PduBB' associates, Alex-488 labeled PduBB' samples are incubated in the presence of salts for 1h at room temperature. 20 µl of protein samples are drop casted onto a glass slide and visualized under the FITC channel using upright fluorescence microscope, model: OLYMPUS BX53 (Olympus, USA).

To check the effect of macromolecular crowding on the self-assembly of PduBB'/PduB'/PduBM38L or PduCDE 2 µl of PEG-6000 (50% w/v stock) is mixed into 20 µl of protein samples in the presence of salts at different concentrations. The samples are drop casted onto a glass slide and visualized under the FITC channel using upright fluorescence microscope, model: OLYMPUS BX53 (Olympus, USA). For co-phase separation experiment, 2 µl of PEG-6000 (50% w/v stock) is added to 20 µl of protein samples (contacting a mixture of Alexa-488 labeled PduBB' and Texas red labeled PduCDE) in the presence of 1 mM MgSO$_4$. The samples are drop casted onto a glass slide and visualized under the FITC and Texas Red channels using upright fluorescence microscope, model: OLYMPUS BX53 (Olympus, USA).

2.2.18 Identification of disordered regions

Percentage of disordered regions and their distribution in shell proteins and enzymes were determined using PONDR web server (http://www.pondr.com/) [73]. The PONDR neural networks use certain sequence attributes like amino acid composition, charge and hydropathy to make predictions. Protein sequences were entered into the space provided in FASTA format and PONDR VSL2 algorithm was used to predict the disordered regions in the protein sequences. PONDR-VLS2 is a highly accurate predictor and has accuracies more than PONDR-VLXT algorithm. Hence, VSL2 was used in the present study to identify disordered

regions in Pdu shell proteins. The raw output file contained percentage of disorder in the proteins and also data points for the plot of disorder propensity versus residue.

The localization of disordered regions and disordered binding regions in the terminal portions of shell protein and enzyme were identified using IUpred 2.0 [74] and ANCHOR webserver (https://iupred2a.elte.hu/) [75]. IUpred 2.0 predicts disordered regions based on the sequence of amino acid and the energy of interactions between the residues. ANCHOR makes use of amino acid sequence as input and identifies those regions which are naturally disordered but may take an ordered conformation upon binding to its partner protein. The advantage of using IUpred and ANCHOR is that it highlights both disordered regions as well as disorder binding regions in the same output file. The protein sequences were entered in the space provided in FASTA format and submitted. The output file contained data points for the plot of disorder/anchor propensity versus protein sequences.

Molecular recognition features among the disordered regions in shell proteins were predicted using MoRFpred webserver (http://biomine.cs.vcu.edu/servers/MoRFpred/) [76]. This webserver helps users to identify residues within a disordered region that may contribute to partner binding by undergoing disorder to order transition. The server uses amino acid sequence of a protein as input and annotations are given based on sequence alignment along with predictions are made using Support Vector Machine (SVM) [76].

2.2.19 Determining residual flexibility

Residual flexibility in proteins was performed using CABS Flex 2.0 web server (http://biocomp.chem.uw.edu.pl/CABSflex2/) [77]. CABS Flex is based on a coarse grained protein modelling tool and is an alternative to 10 ns all atom molecular dynamics simulation. It can generate near-native protein dynamics at a significantly reduced computational cost.

Since the crystal structure of PduB and PduN are not solved, we performed homology modelling of the monomeric unit of PduB and PduN using Galaxy TBM server [69] and the structures obtained were saved as PDB files. The PDB files were uploaded into the CABS Flex web server. The simulations are run for 50 cycles and cycles between the trajectory frames is set at 50. This allows a total of 1000 models to be generated in the simulation trajectory, out of which every 50th model is saved. Out of all the saved models, we select the most preferred model displayed in the result for further analysis.

2.2.20 Limited proteolysis

20 µl of 0.5 mg/ml of shell proteins (PduBB'/PduB/PduBM38L) was treated with trypsin (at shell protein: enzyme ratio of 100:1 w/w) for different time periods. The reaction was quenched by adding 1X SDS running buffer and boiling the samples at 95°C for 10 min. The protein samples were visualized by running on 12% SDS PAGE gel.

2.2.21 Docking and MD simulation

The N-terminal segments of PduB and PduD were modeled using Galaxy TBM server. Peptide-peptide docking between modeled N-terminals of PduB and PduD was performed using CABS dock server [78]. The monomeric units of PduB and PduD/ΔN PduD were modeled using Galaxy TBM server. Molecular docking of PduB-PaduD and PduB-ΔN PduD was carried out using ClusPro server. The docked models were then subjected to MD simulation using the QwikMD plugin of VMD and run using NAMD version 2.12. The protein structures are solvated inside a box with its edge 15Å from the protein molecule and Na+ and Cl- ions are added at a concentration of 150 mM. Energy minimization is carried out for 50000 steps followed by conventional MD simulation for 20 ns.

Chapter 3

Chaperone Like Activity of a Major Shell Protein

Enzyme (PduCDE)

Heat Shock

Denaturation and aggregation of enzyme

Shell protein (PduBB')

Shell protein:Enzyme (PduBB':PduCDE)

Heat Shock

Stabilization of enzyme by shell protein

Scatter

PduCDE

PduBB':PduCDE

Time

Activity

PduBB':PduCDE

PduCDE
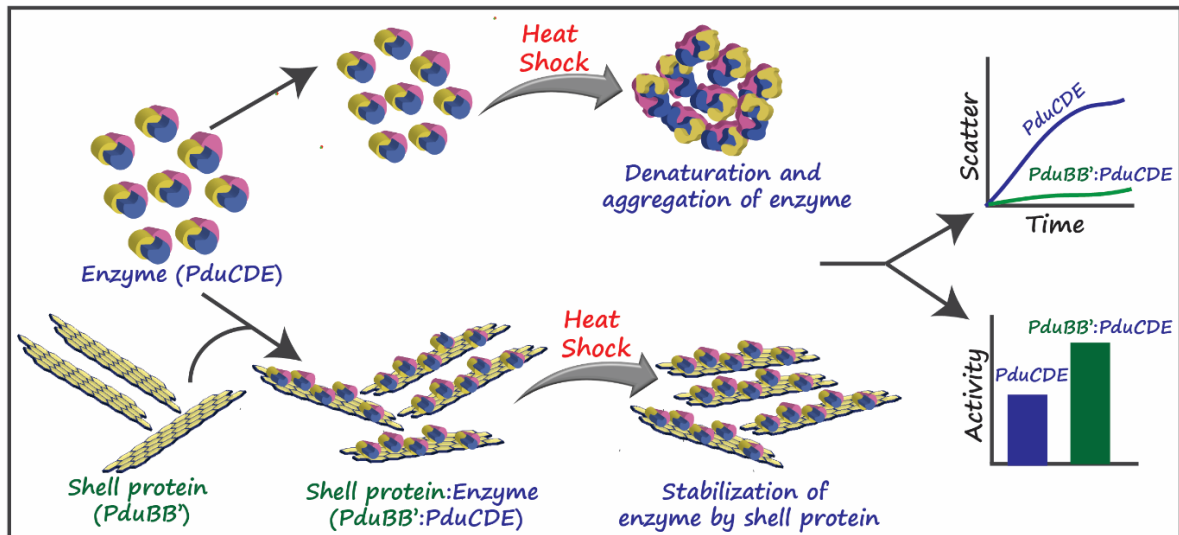
3.1 Introduction

The first chapter gave us a detailed information about how certain bacterial species produce proteinaceous compartments called bacterial microcompartments (MCP) to aid metabolism that help them to survive under nutrient specific conditions.[13, 16, 79] Studying the structure and function of such sub-cellular organizations may not just provide us with vital information regarding pre-biological compartmentalization [80] but will also pave the way for the development of novel synthetic reactors for future applications.[81] [82] We now try to understand how compartmentalization is important for enzyme function and investigate if shell proteins beside regulating metabolite transport have additional functions Among all the MCPs discovered so far, 1,2-propanediol (1,2-PD) utilization microcompartment (PduMCP) is the most complex yet most studied (Fig. 1).[56, 83-86] In this chapter we use 1,2-propanediol utilization microcompartment (PduMCP) as a model to study the effect of molecular confinement on the stability and catalytic activity of native enzymes in microcompartment. Although these studies have generated a lot of information regarding the mechanism by which the PduMCP functions, very little is known about the impact of the shell proteins on the catalytic efficiency and stability of the enzymes in a confined environment of PduMCP. Since PduMCP is made up of several copies of multiple shell proteins and enzymes, we have taken a minimalistic approach to understand how a particular shell protein would influence the catalytic efficiency of a native enzyme of PduMCP, under thermal stress. In this context, we explored the effect of the major shell protein PduBB' on the stability and activity of the signature enzyme diol dehydratase PduCDE using a combination of biophysical and biochemical techniques [50].

The shell protein PduBB' is expressed as a combination of PduB (270 amino acids, 28 kDa) and PduB' (233 amino acids, 25 kDa) **(Fig. 3.1, a-b)**, from alternative start sites but the same end codon in the *pdu operon* [84],[87]. The translated protein pair PduBB' comprises a major portion of PduMCP outer shell and deletion of this pair impairs the PduMCP formation.[46] The extended N-terminal region in PduBB' has been suggested to link the shell proteins of PduMCP with its encapsulated enzymes.[47] However, the significance of this combination in the context of PduMCP stability and shell-enzyme interaction is not known. In this chapter we aim at understanding the importance of PduB and PduB' combination in the PduMCP stability and activity. We observe that PduBB' preserves the catalytic efficiency of PduCDE under thermal stress and prevents temperature-induced unfolding and aggregation of PduCDE *in vitro*. While both PduB and PduB' interact with the enzyme with micromolar affinity, only

the PduBB' combination influences its activity and stability, highlighting the importance of the unique PduBB' combination in the functioning of PduMCP.
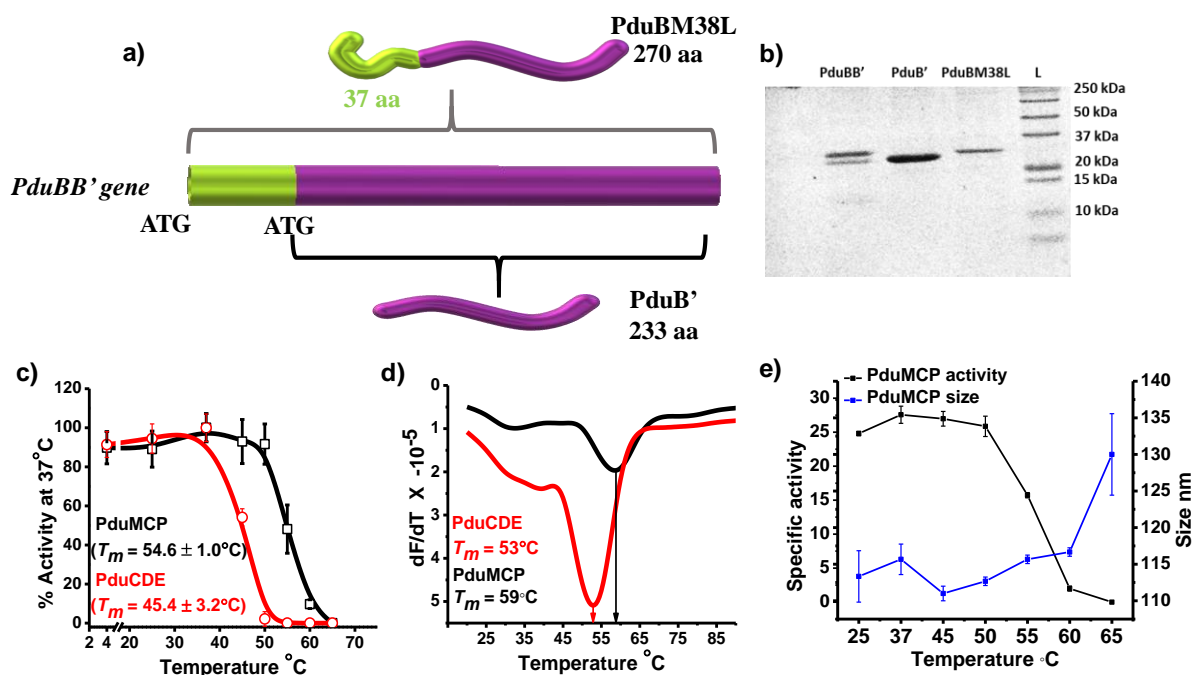
Figure 3.1: a) Schematic representation of *PduBB' gene* and its protein products PduB and PduB', b) 12% SDS PAGE of shell proteins (i) PduBB', PduB', PduB (generated as single mutant PduBM38L), c) Temperature dependant diol dehydratase assay using PduMCP (black) and PduCDE (red), (b) Derivative plot of change in intrinsic protein fluorescence (at 328 nm) as a function of temperature for PduMCP (black) and PduCDE (red), d) Specific activity and hydrodynamic diameter of PduMCP at temperatures 25°C, 37°C, 45°C, 50°C, 55°C, 60°C, and 65°C.

3.2 Diol dehydratase catalytic activity of encapsulated PduCDE is conserved under thermal stress

To understand the effect of encapsulation on the functioning of native enzyme in PduMCP, the diol dehydratase activity of PduMCP and purified PduCDE is compared post thermal shock. PduCDE or PduMCPs samples are incubated at different temperatures in the range of 4°C-60°C for 2 min, followed by equilibration at 37°C prior to performing diol dehydratase assay at 37°C (**Fig. 3.1c, Table. 3.1).** This method allows us to assess how efficiently the two systems (free and encapsulated enzyme) retains their optimum activity after a thermal shock.

PduMCP shows maximum activity at 37°C which is retained till 50°C, after which a gradual decrease is observed. PduMCP shows $A_{50\%}$ (50% of the maximum activity at 37°C) at ~55 °C. Beyond 60°C, PduMCP shows no activity. In contrast, the free enzyme PduCDE starts to lose its activity immediately after 37°C and shows $A_{50\%}$ at ~45°C. No activity of the bare enzyme is detected beyond 50°C. Next the conformational stability of the bare enzyme PduCDE and PduMCP is compared using intrinsic protein fluorescence. An amino acid analysis reveals that the PduMCP enzymes have a higher content of aromatic residues than the shell proteins **(Table. 3.2 and 3.3)**. This implies that the encapsulated enzymes would contribute the most to the intrinsic protein fluorescence from PduMCP, allowing a direct comparison of the properties of the bare and encapsulated enzymes. The derivative plot of fluorescence intensity (at 328 nm) with respect to temperature for PduMCP and PduCDE is shown in **Fig. 3.1d**. The T_m of PduMCP (59°C) is found to be higher than that of PduCDE (T_m = 53°C) indicating a higher conformational stability of the encapsulated enzyme within the PduMCP compared to the free form. When probed by dynamic light scattering (DLS), PduMCP retains an average hydrodynamic diameter of 116 ± 1 nm till 60°C **(Fig. 3.1e)**. Similar trends in size variation have been reported earlier by electron microscopy.[88] The findings in this section leads us to two key conclusions. First, the encapsulation of the enzyme PduCDE within PduMCP gives an advantage of ~10°C in the activity regime of the enzyme. Second, PduMCP remains structurally intact up to 60°C, however PduCDE encapsulated within, starts to lose its activity beyond 50°C **(Fig. 3.1e)**. After 60°C, when the outer shell begins to deform, enzyme activity is completely lost, implying that shell proteins play a critical role in preserving enzyme activity at higher temperatures. In the next section we determine the role of major shell protein PduBB' in chaperoning PduCDE under thermal stress.

Table 3.1: Specific Activity of diol dehydratase in encapsulated PduMCP and in free form as PduCDE		
Temperature °C	**PduMCP**	**PduCDE**
	Activity μmol mg^{-1} min^{-1}	**Activity μmol mg^{-1} min^{-1}**
4	**22.9±1.2**	**21.3±2.0**
25	**22.7±2.3**	**22.0±1.7**
37	**25.5±1.8**	**23.3±1.6**
45	**23.7±2.8**	**12.6±1.0**
50	**23.3±2.6**	**0.50±0.8**
55	**12.2±3.1**	**0**
60	**2.4±0.6**	**0**

Table 3.2: Number of aromatic amino acids in Pdu shell proteins		
Shell proteins	No. of tryptophan residues	No. of tyrosine residues
PduA	0	1
PduB	0	6
PduJ	0	1
PduM	0	1
PduN	0	0
PduK	2	0
PduT	2	2
PduU	0	2

Table 3.3: Number of aromatic amino acids in Pdu enzymes		
Enzymes	No. of tryptophan residues	No. of tyrosine residues
PduCDE	5	24
PduP	1	5
PduL	0	2
PduQ	3	5
PduW	3	9
PduG	1	8
PduS	3	6
PduH	2	1

3.3 Shell protein PduBB' shows chaperone like behavior towards PduCDE

The shell protein PduBB' when overexpressed shows 2D sheets as observed under TEM (**Fig. 3.2a, i and ii**). Heterotrimeric PduCDE (**Fig. 3.2b**) forms assemblies as well, but unlike shell protein, PduCDE assemblies lacks a distinct sheet like morphology (**Fig. 3.2a, vii and viii**). Temperature dependent structural studies using circular dichroism of PduBB' and PduCDE show loss of their native conformation (**Fig. 3.2, c-d**) around ~80°C and ~50°C respectively (**Fig. 3.2e**), suggesting higher thermal stability of shell protein PduBB' compared to enzyme PduCDE. This observation coupled with the observation made in previous section prompted us to test if the thermostable shell protein PduBB' could also influence the stability of the enzyme PduCDE under thermal stress. The thermal aggregation of PduCDE in the absence and presence of PduBB' is probed using Rayleigh scattering. An increase in the scattering at 360 nm is observed in PduCDE upon increasing the temperature to 45°C indicating aggregation of the enzyme (**Fig. 3.2f**). However, we did not see any increase in scattering intensity of PduCDE in the presence of the shell protein PduBB'. A PduBB': PduCDE molar ratio of 8:1 is required to completely prevent thermal denaturation and aggregation of the enzyme at 45°C (**Fig. 3.2f**), indicating that the shell protein PduBB' protects PduCDE from

thermal aggregation. Since PduBB' is a combination of PduB and PduB', we next question their individual role in chaperoning the thermal aggregation of the enzyme. PduB is generated as a single mutant clone of PduBB' (PduBM38L) **(Fig. 3.1b)** (PduB and PduBM38L will be used inter-changeably in this thesis henceforth). Both PduB' and PduBM38L individually fail to prevent the aggregation of PduCDE at shell protein: enzyme ratio of 8:1 **(Fig. 3.2g)**. This observation underscores for the first time the significance of the PduBB' combination in the context of enzyme stability. We further observe that physical mixture of PduB'+PduBM38L does not protect PduCDE from thermal denaturation suggesting that mere mixing of PduB' and PduBM38L cannot mimic the properties of the co-expressed combination PduBB' **(Fig. 3.2g)**. These observations show unique property of PduBB' to prevent thermal denaturation of the native enzyme PduCDE.

3.4 Enhanced activity of PduCDE in the presence of the shell protein PduBB'

The activity of PduCDE is checked in the absence and presence of PduBB' under thermal stress. At PduBB': PduCDE molar ratio of 8:1, PduCDE shows ~56% higher specific activity than the bare enzyme, post thermal shock at 45°C **(Fig. 3.3a)**. Post heat shock at 50°C, the activity of the bare enzyme PduCDE is completely lost. However, we do notice an activity of ~3 μmol min^{-1} mg^{-1} in the presence of PduBB'. To check if this enhancement in activity of PduCDE is due to specific interactions between the enzyme and PduBB', the experiment is repeated using PduB' and PduBM38L. We find no difference in PduCDE activity in the absence or presence of PduB' or PduBM38L (at 8:1 molar ratio) post thermal stress. Next, we investigate whether a physical mixture of PduBB' and PduCDE could affect the enzyme's catalytic activity in the absence of thermal stress. At lower PduBB' to PduCDE molar ratio, no significant difference in enzyme activity is observed. However, increasing the molar ratio to 64:1 or higher results in a 33% increase in PduCDE activity **(Fig. 3.3b)**. At PduBB': PduCDE molar ratio of 64:1, PduCDE shows ~78% higher activity than bare PduCDE post heat shock at 45°C, suggesting a protective role of PduBB' **(Fig. 3.3c)**. Post heat shock at 50°C, no PduCDE activity is detected in the absence of PduBB', whereas in the presence of PduBB', PduCDE shows an activity of 4.2 ± 0.6 μmol min^{-1} mg^{-1} **(Fig. 3.3c)**. This experiment demonstrates that the shell protein PduBB' not only increases the catalytic activity of PduCDE but also protects it from thermal stress. Under similar conditions, the constituent shell proteins of PduBB' (PduB' and PduBM38L) show no effect on the catalytic activity of PduCDE **(Fig. 3.3, d-e)**. This reinforces the idea that the presence of both PduB and PduB' is critical for the protective role of PduBB'.
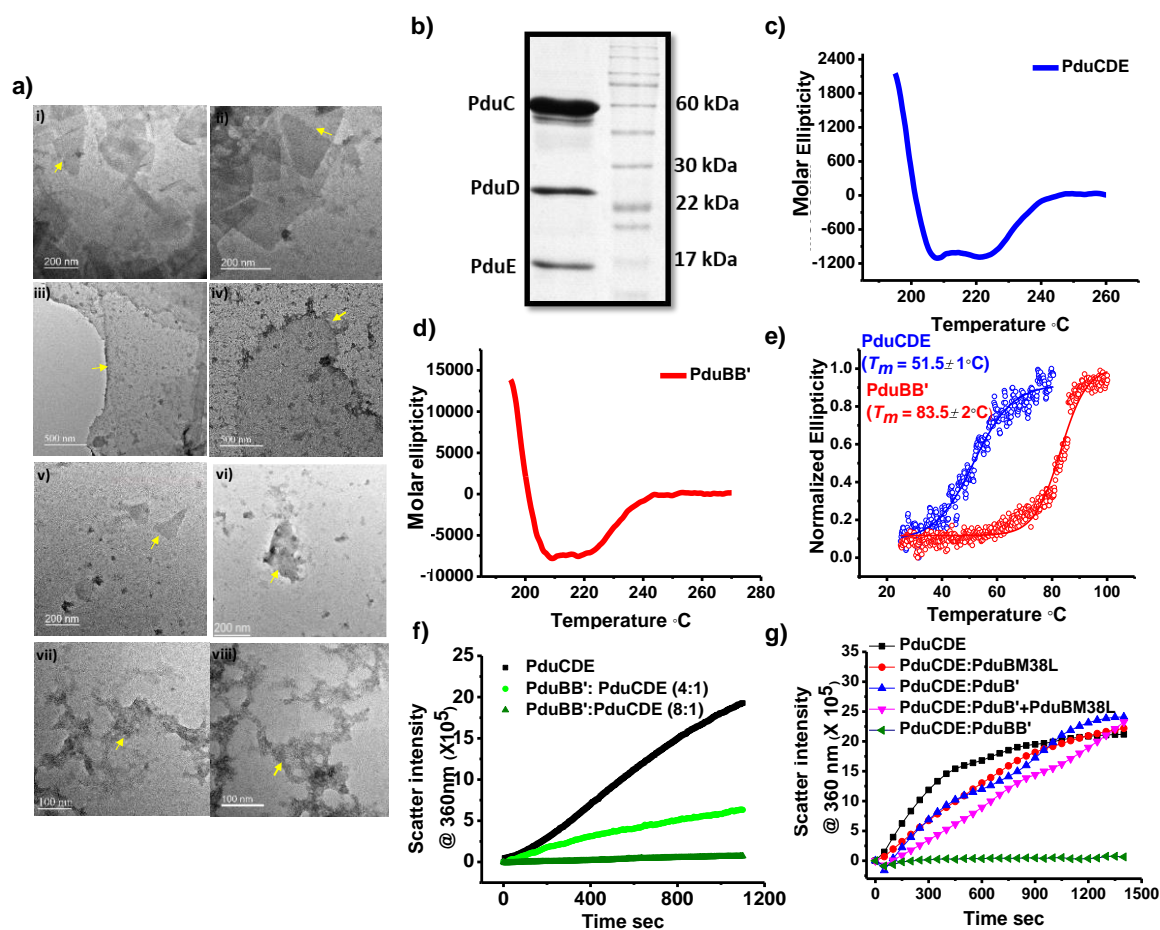
Figure 3.2: a) TEM image of shell protein PduBB' at concentration of 0.2 mg/ml (i) and (ii), TEM image of PduB' (0.2 mg/ml) (iii) and (iv), TEM image of PduBM38L (0.2 mg/ml) (v) and (vi), TEM image of PduCDE at 0.2 mg/ml (vii) and (viii), b) SDS PAGE of PduCDE, c) CD spectra of PduBB', d) CD spectra of PduBB', e) CD melt spectra of PduBB' (red) and PduCDE (blue), the plot shows normalized molar ellipticity measured at 222 nm at different temperatures f) Normalized scatter intensity measured at 360 nm for PduCDE alone and PduCDE in the presence of PduBB' mixed in different shell protein: enzyme molar ratio (g) Normalized scatter intensity measured at 360 nm for PduCDE alone (black), PduCDE in the presence of PduBB' (green), PduB' (blue), PduBM38L (red), PduB'+PduBM38L (pink) at shell protein:enzyme molar ratio of 8:1 and exposure to a temperature of 45°C.

Box 3.1 Aggregation kinetics and enzyme assay using non-native protein

a) **Aggregation kinetics measured at 360 nm for PduCDE alone and PduCDE in the presence of lysozyme at lysozyme: PduCDE molar ratio of 16:1, b)** Specific activity of alcohol dehydrogenase (Alcdh) (μmol min^{-1} mg^{-1}) in the absence and presence of shell protein PduBB' (64:1 molar ratio), post exposure to temperatures 37°C and 60°C.

These control experiments using non-native proteins negate the possibility of non-specific protein-protein interactions in our observations

The enzyme assays and aggregation kinetics together suggest that a shell-to-enzyme molar ratio of 8:1 for is sufficient to provide thermal stability to the enzyme PduCDE. However, a high shell-to-enzyme ratio is required to observe the influence of shell protein PduBB' on PduCDE activity under optimal conditions **(Fig. 3.3b)**. The requirement for a high shell protein to enzyme ratio may be explained by the fact that the experiment depicted in **Fig. 3.3b** is a bulk ensemble system and does not represent a volume confined environment of PduMCP. In a confined environment like PduMCP, the overall activity of enzyme is higher than the enzyme in free solution. This is evident from the comparative specific activity of PduMCP and PduCDE shown in **Table. 3.1**. At 37°C, the specific activity values of PduCDE and PduMCP are similar. However, to get the specific activity values the absolute activity values of PduMCP and PduCDE are divided by the amount of PduMCP or PduCDE present in the reaction mixture. Since PduMCP is a multi-protein complex and the enzyme PduCDE is encapsulated in a lower amount, the absolute activity value of encapsulated PduCDE would be much higher than the PduCDE present in free solution. This suggests that limiting enzymes' space and volume improves their performance. It would be difficult to mimic such environment by taking one shell protein and enzyme pair in free solution. Therefore, a high shell-to-enzyme would be necessary to influence the activity of the enzyme under optimum condition *in vitro*

Mixing PduBM38L and PduB' in different molar ratios also has no effect on the catalytic activity of PduCDE **(Fig. 3.3f)**, proving that a physical mixture of the two component shell

proteins cannot mimic the properties of PduBB'. This raises a question of why PduBB', rather than its individual components, improves the catalytic efficiency and protects the enzyme from thermal stress. We address this problem by taking the structural stability, self-assembly, solubility, and affinity for PduCDE of the three PduBB' variants into account.
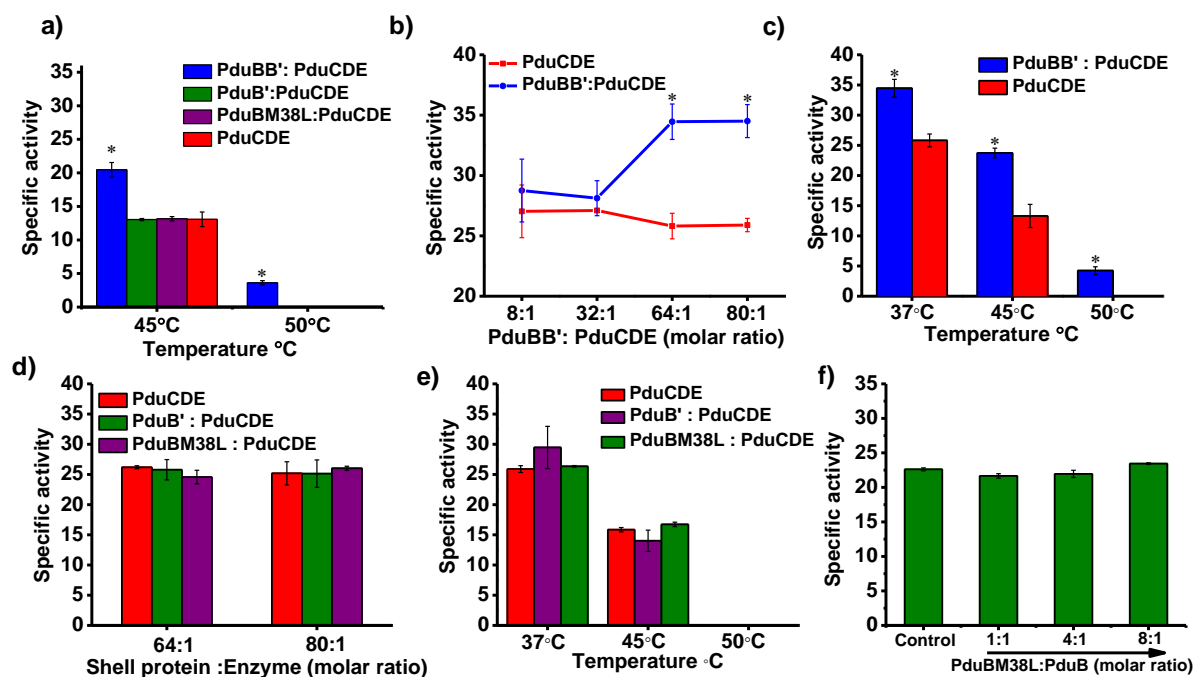
Figure 3.3: a) Specific activity (µmol min^{-1} mg^{-1}) of PduCDE post heat shock in the absence and presence of Shell proteins mixed in shell protein: enzyme of 8:1 molar ratio, b) Specific activity (µmol min^{-1} mg^{-1}) of PduCDE (without thermal stress) in the absence and presence of PduBB' at different shell protein: enzyme molar ratio, c) Specific activity (µmol min^{-1} mg^{-1}) of PduCDE in the absence and presence of PduBB' at PduBB': PduCDE molar ratio of 64:1 post exposure to 37°C, 45°C and 50°C, d) Specific activity of PduCDE (µmol min^{-1} mg^{-1}) in the absence and presence of PduB' and PduBM38L at different shell: enzyme molar ratio, e) Specific activity (µmol min^{-1} mg^{-1}) of PduCDE in the absence and presence of PduB' and PduBM38L at shell: enzyme molar ratio of 64:1, post heat shock at 37°C, 45°C and 50°C, [*p ≤ 0.05]. The data are mean of the three independent experiments, f) Specific activity of PduCDE (µmol min^{-1} mg^{-1}) in the absence (control sample) and presence of physical mixture of PduBM38L and PduB',

3.5 Self-assembly of PduB and PduB' leads to the formation of stable and soluble shell protein PduBB' combination *in vitro*

As seen in TEM, all three PduBB' variants demonstrate the ability to self-assemble and form sheets **(Fig. 3.2a, i-vi)**. However, careful examination of their morphologies reveals a distinct difference in the self-assembly behavior of the three PduBB' forms. In contrast to PduBB', which forms multiple discrete sheets **(Fig. 3.2a, i-ii)**, PduB' forms large extended sheets, indicating a higher tendency for PduB' to self-assemble **(Fig. 3.2a, iii-iv)**. In addition, unlike PduBB', very few sparsely settled and relatively small sheets and assemblies of PduBM38L are visible under TEM **(Fig. 3.2a, v-vi)**. Among the three shell protein variants, PduBM38L shows the least self-assembly characteristics.

In solution, PduB' self-associates to give a turbid solution with high optical density and became soluble only upon heating at 60°C **(Fig. 3.4a)**. The turbidity is not observed in case of PduBM38L or for the co-expressed combination PduBB'. Heating PduB' at 60°C does not alter the secondary structure **(Fig. 3.4b)** or sheet forming property of the PduB' **(Fig. 3.4b, inset)**. An important difference between PduB and PduB', is the presence of extended 37 amino acids N-terminal in PduB **(Fig. 3.4c)**, and probably this assist in providing solubility to the shell protein PduBB' or PduBM38L, thus regulating their self-association. Furthermore, temperature dependent CD spectra of shell proteins reveals lower structural stability of PduBM38L as compared to PduBB' and PduB' **(Fig. 3.4, e-g).** Both PduBB' and PduB' retain their secondary structures up to 80°C and show loss in their native conformation further **(Fig. 3.4, e-f)**. PduBM38L shows a transition around 60°C indicating its lower thermal stability as compared to PduBB' and PduB **(Fig. 3.4g)**. All the three shell proteins regain their native conformation upon cooling post thermal melt **(Fig. 3.4, h-j).** Lower ellipticity of PduB' pre melt **(Fig. 3.4i)** may be attributed to the formation of PduB' associates in solution, displaying high optical density at 600 nm **(Fig. 3.4a)**. Dissociation of PduB' associates due to heat results in increased ellipticity post melt.

This suggests that presence of both PduB and PduB' is necessary for the formation of a soluble and stable shell protein combination PduBB' *in vitro*.

The self-assembly behavior of the three variants of PduBB' is probed by monitoring their size distribution in solution phase at different shell protein concentrations using dynamic light scattering (DLS) **(Fig. 3.5)**. PduB' shows a size distribution in micron range at 0.1 mg/ml and shows no significant change upon increasing the protein concentration to 0.8 mg/ml **(Fig.**

3.5a). PduBM38L shows size distribution between 10 nm to 300 nm at all concentrations of shell protein ranging from 0.1-0.8 mg/ml **(Fig. 3.5b)**, suggesting a lower tendency to self-associate as compared to PduB'. PduBB' displays a distinct self-assembly behavior in solution. At low shell protein concentration of 0.1 mg/ml, it shows a size distribution between 10 nm to 500 nm. With increasing shell protein concentration to 0.8 mg/ml, PduBB' undergoes self-association giving 50 % of the size distribution in the micron range and rest, less than 100 nm **(Fig. 3.5c)**. The TEM images, turbidity assay, and DLS study together indicate that the tendency to self-assemble is highest in PduB' and lowest in PduBM38L. Circular dichroism studies show that all the three shell proteins retain their secondary structures at all concentrations in the range tested **(Fig. 3.5, d-f)**. Noticeably, the ellipticity of PduBB' decreases as the concentration of shell protein is increased **(Fig. 3.6f)**. This decrease in ellipticity may be due to the formation of higher order assemblies (DLS results shown in **Fig 3.5c**) at higher concentrations leading to structural alterations. An auto-optimized self-association in case of PduBB', suggests a vital role of N-terminal region in regulating the self-assembly of PduBB' post translation which perhaps has implications in tuning the size of the mature PduMCP.
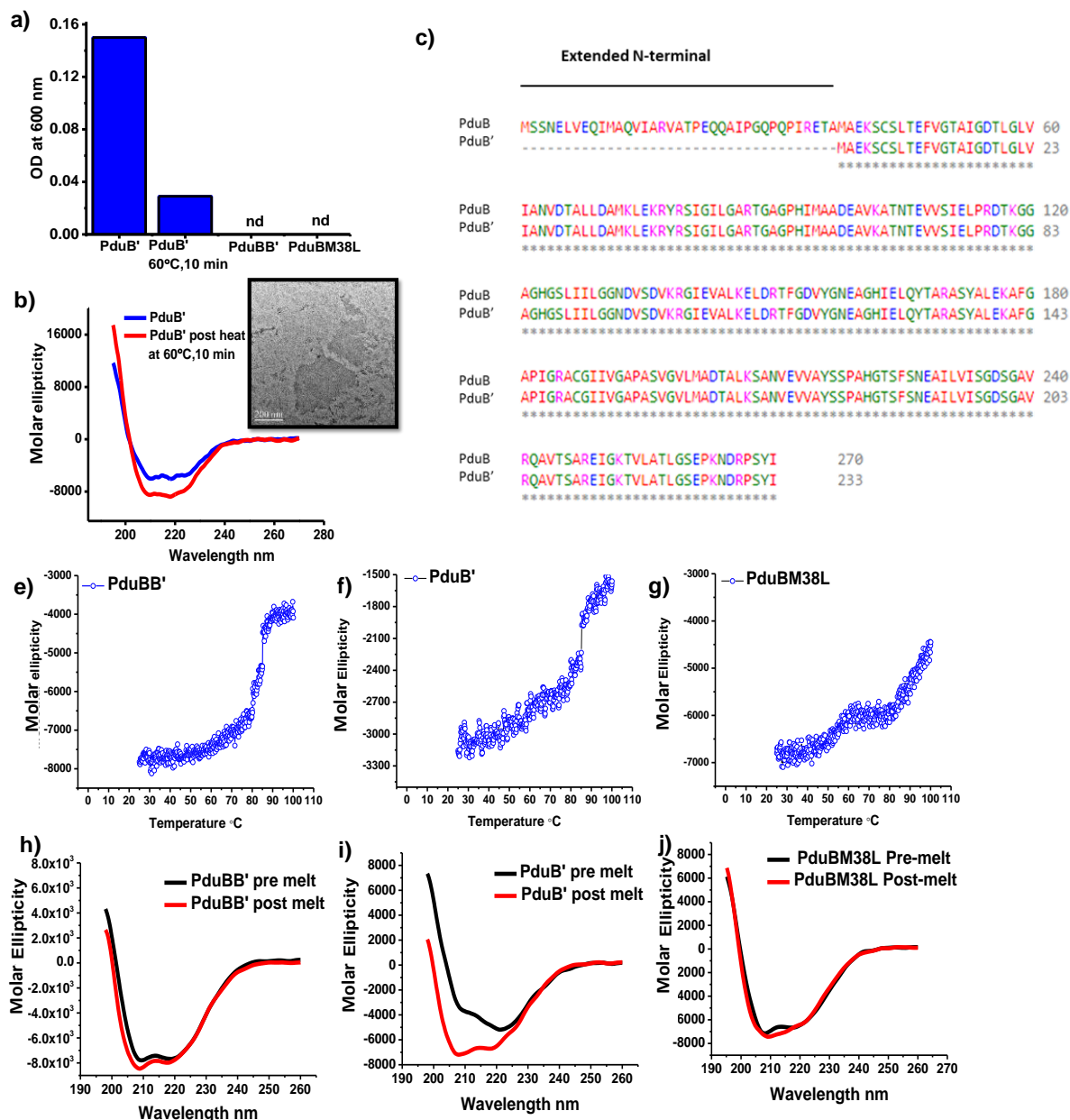
Fig. 3.4: a) Turbidity assay: Scattering at 600 nm for PduB', PduBB' and PduBM38L (0.25 mg/ml each), b) CD spectra of PduB' (0.25mg/ml) pre and post heating at 60°C for 10 min, TEM image of PduB' post heating at 60°C for 10 min (in inset), c) Amino acid sequence of PduB and PduB'. PduB (270 residues) is longer than PduB' (233 residues) and has an extended 37 amino acids N-terminal region. e-g) Plot of Ellipticity measured at 222 nm versus temperature for (e) PduBB', (f) PduB', (g) PduBM38L. h-j) Pre and post melt spectra of the shell proteins (h) PduBB', (i) PduB', (j) PduBM38L.
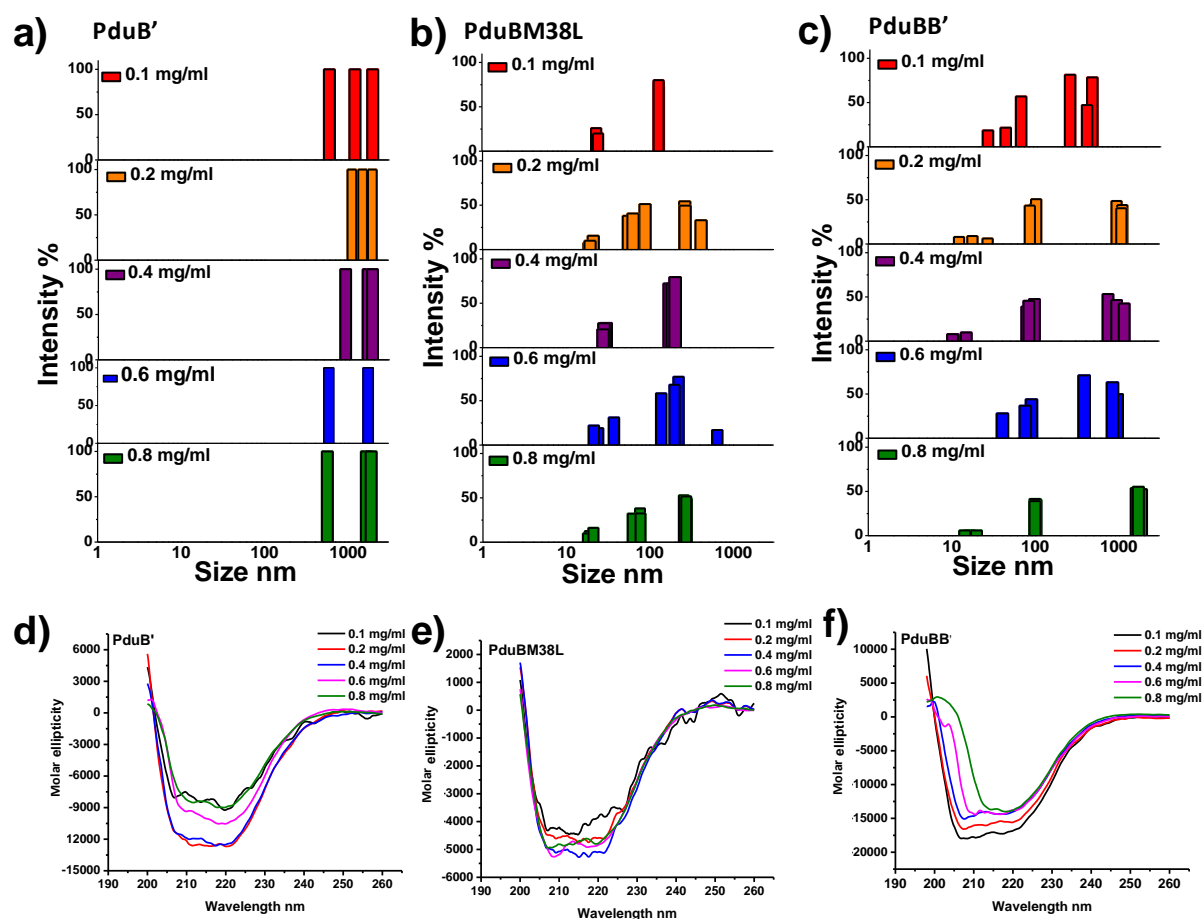
Fig. 3.5: Concentration dependent size distribution of shell protein PduBB' and its variants PduB' and PduBM38L, measured using dynamic light scattering. Intensity percentage showing size distribution of PduB' (a), PduBM38L (b) and PduBB' (c) in solution at protein concentrations ranging from 0.1-0.8 mg/ml. (d-f) Concentration dependent CD spectra of (a) PduBB', (b) PduB' (c) PduBM38L

[Note: For each concentration, three readings in DLS are recorded. Each reading gives three values that add up to 100 %. In case of PduB' however, due to the formation of large micron size associates, for every reading we get 100% intensity in the micron range. Each bar shows 100% intensity value because for all the three readings we get a single size in micron range with 100% intensity].

3.6 Probing the affinity of PduCDE towards PduBB'

The interaction of the enzyme PduCDE with the variants of PduBB' is probed using biolayer interferometry (BLI) **(Fig 3.6, a-d)**. PduCDE is immobilized onto optical sensors **(Fig. 3.6a)** and association and dissociation kinetics of shell proteins is measured. PduBB' associates with PduCDE with dissociation constant (K_d) of $(7.41 \pm 0.045) \times 10^{-7}$ M **(Fig. 3.6b)**. For

PduB', as shell protein concentration increases, the association response gets saturated, followed by faster decay kinetics **(Fig. 3.6c)**. This could be due to weaker interactions of PduB' towards PduCDE with K_d value of $(1.36\pm0.007) \times10^{-6}$ M. PduBM38L shows a similar trend, indicating weaker interactions with PduCDE with K_d value of $(3.31\pm0.017) \times10^{-6}$ M **(Fig. 3.6d)**. This experiment reveals that PduBB' has greater affinity for PduCDE than its individual components (PduB' and PduBM38L). A higher affinity of PduCDE towards PduBB' may contribute to a better scaffolding of the enzyme by the shell protein which in turn would confer stability to the enzyme under thermal stress. Control experiments with non-specific proteins further eliminate the possibility of non-specific interactions in our experiments **(Fig. 3.6, e-f)**.

The dynamics of shell proteins in the vicinity of enzyme in solution is studied using fluorescence anisotropy. The presence of cysteine residues in PduBB' (Cys43 and Cys 187) **(Fig. 3.7a)** allows us to label the shell protein with thiol reactive acrylodan dye. We observe no change in the size distribution or secondary structure of PduBB' after dye labelling. **(Fig. 3.7, b-c)**. Time resolved fluorescence anisotropy of shell proteins (PduBB'/PduB'/PduBM38L) is carried out in the absence and presence of PduCDE mixed in 1:1 molar ratio **(Fig. 3.7, d-f)**. The decay kinetics are best fitted to two exponential function (*chapter 2*, *Eq. 1*) which provides two rotational correlation times **(Table. 3.4)**. We get a shorter rotational correlation time in sub-nanoseconds (0.2-0.5 ns) in case of all the three shell proteins (both in the absence and presence of PduCDE). The sub-nanosecond correlation time corresponds to the tumbling motion of acrylodan dye molecule **(Fig. 3.7g)**. The second and the longer correlation time seen in case of all the three shell proteins corresponds to the regional dynamics within the shell proteins. In case of both PduBB' and PduB', an increase in the longer correlation time in the presence of PduCDE is observed **(Fig. 3.7, d-e, Table. 3.4)** which implies interaction between the shell proteins and enzyme. We do not see any difference in the decay kinetics of PduBM38L in the absence and presence of PduCDE at equimolar ratio **(Fig. 3.7f)** and increase in the longer correlation time is observed only when enzyme and shell protein are mixed in higher molar ratio (4:1) **(Fig. 3.7h, Table. 3.4)**.
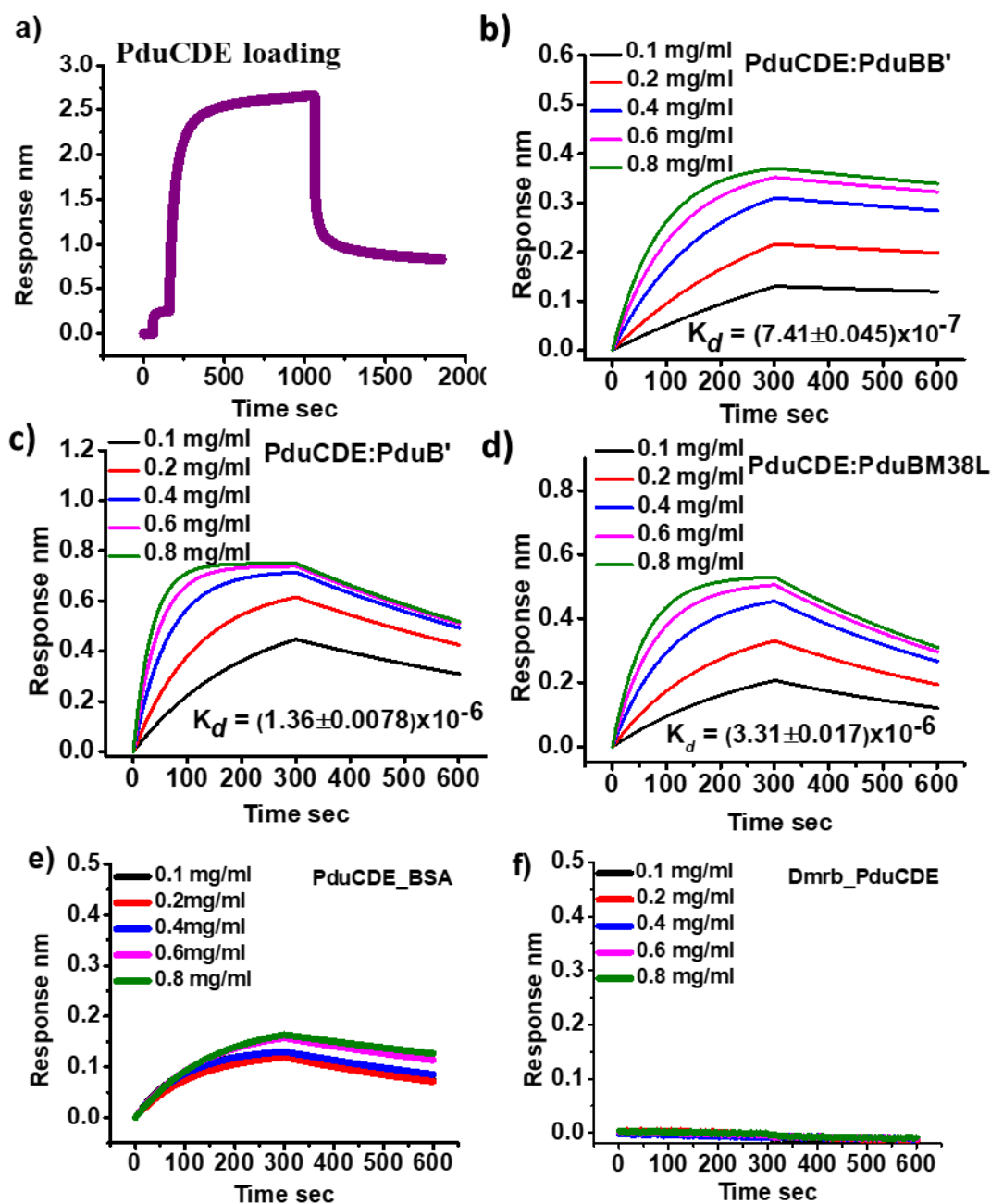
Fig. 3.6: a) Loading of PduCDE using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC)/N-hydroxysuccinimide (NHS) method onto the amine reactive second generation (AR2G) biosensors, b-d) Interactions between PduCDE and the shell proteins (b) PduBB', (c) PduB' and (d) PduBM38L, using biolayer interferometry. e-f) Interactions between PduCDE and e) BSA, f) His-tagged Dmrb (Dihydromethanopterin reductase) protein.

Note: Saturated response in the association kinetics between PduCDE and BSA shows non-specific interactions between the two. No association response observed between PduCDE and His-tagged Dmrb protein, eliminating the role of His tag in association between the PduCDE and shell proteins

One possible explanation for such an observation is that PduBM38L has a lower affinity for PduCDE than PduBB'. Because of the lower binding affinity, a higher concentration of PduCDE would be required to affect the anisotropy decay of PduBM38L. However, despite having a binding affinity similar to that of PduBM38L, PduB' exhibits an increase in rotational correlation time at a 1:1 molar ratio. This may be due to the propensity of PduB'for self-association in solution **(Fig. 3.4a)**. As a result, even if fewer PduCDE molecules bind to the PduB' (due to lower binding affinity), the resulting PduB': PduCDE assemblage would be heavier and move slower, increasing the rotational correlation time. Further, the concentration of shell proteins used in the experiment could also influence the anisotropy experiment's outcome. Shell proteins are used at 3 µM concentrations for anisotropy studies, which is close to the K_d value of the PduBM38L: PduCDE interaction. This may explain why the rotational correlation time increases only when the PduCDE: PduBM38L ratio is increased to 4:1. Together, our results show that all the three shell proteins PduBB', PduB' and PduBM38L have the ability to interact with PduCDE. However, PduBB' shows highest affinity towards PduCDE **(Fig. 3.6b).** Higher binding affinity and a naturally optimized self-assembly behavior of PduBB' **(Fig. 3.5c),** would provide a suitable scaffolding necessary for the improved catalytic activity and enhanced thermal stability of the enzyme.
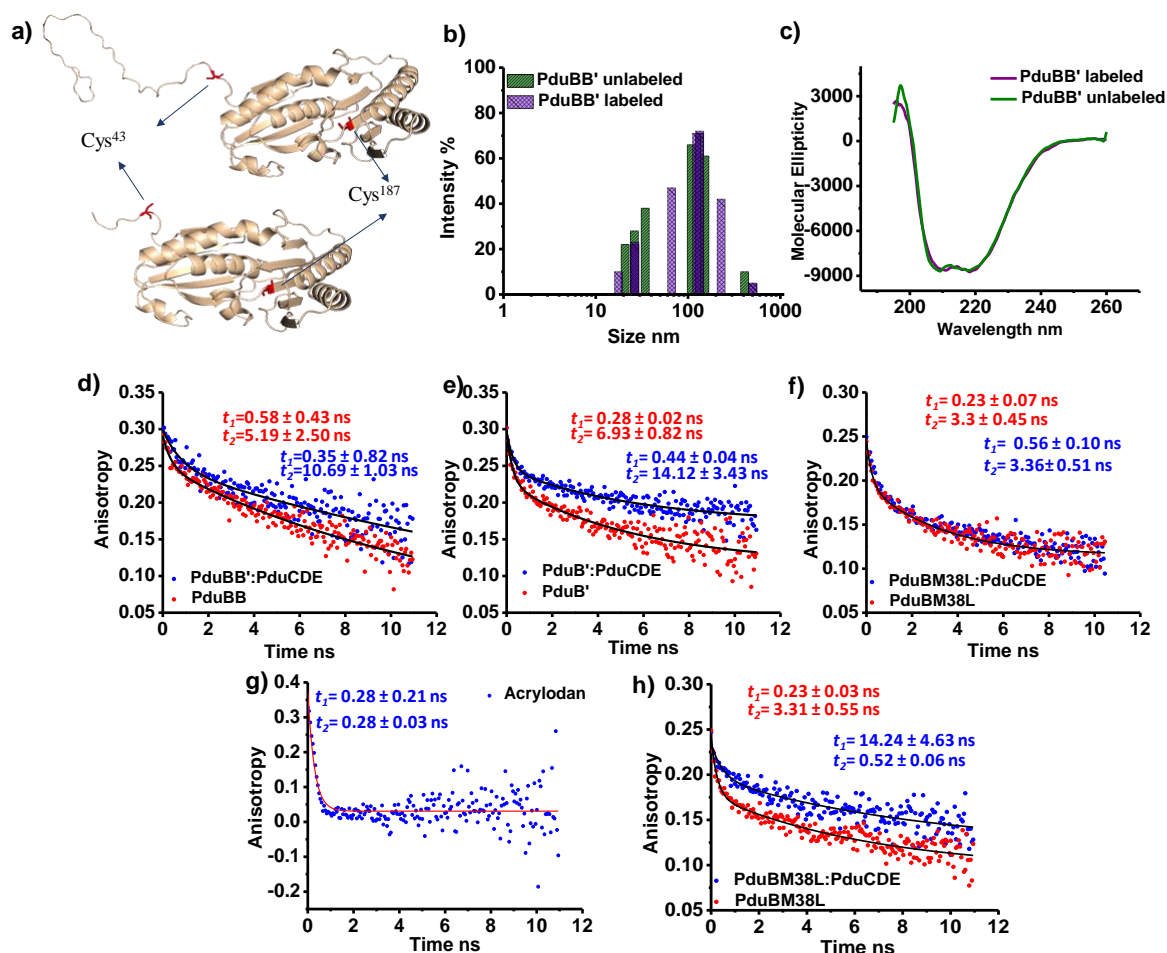
Fig. 3.7: a) Homology modeling of PduBB' from Salmonella enterica using AIDA server (https://aida.godziklab.org/). Cysteine residues have been marked in red. b) Size distribution of unlabeled and Acrylodan labeled PduBB', c) CD spectra of PduBB' before and after fluorophore labeling, d) Time resolved anisotropy decay of shell protein PduBB' in the absence and presence of enzyme PduCDE mixed in 1:1 molar ratio, e) Time resolved anisotropy decay of shell protein PduB' in the absence and presence of enzyme PduCDE mixed in 1:1 molar ratio. f) Time resolved anisotropy decay of PduBM38L in the absence and presence of the enzyme PduCDE mixed in 1:1 molar ratio. g) Time resolved anisotropy decay of free acrylodan dye, best fitted to bi-exponential model. Free acrylodan gives rotational correlation time in sub-nano second range. h) Time resolved anisotropy decay of PduBM38L in the absence and presence of the enzyme PduCDE mixed in enzyme: shell of 4:1 molar ratio.

Table 3.4: Rotational correlational time and their corresponding amplitude obtained using bi-exponential decay model.				
	Rotational correlation time	Amplitude	Rotational correlation time	Amplitude
	- PduCDE		+ PduCDE	
PduBB'	0.58 ± 0.43 ns	3.79	0.35 ± 0.82 ns	1.09
	5.19 ± 2.50 ns	96.21	10.69 ± 1.03 ns	98.91
PduB'	0.28 ± 0.02 ns	7.73	0.44 ± 0.04 ns	2.25
	6.93 ± 0.82 ns	92.27	14.12 ± 3.43 ns	97.75
PduBM38L	0.23 ± 0.03 ns	9.62	0.52 ± 0.06 ns	4.53
	3.31 ± 0.55 ns	90.38	14.24 ± 4.63 ns	95.47

Using thermal shift assay we also show that the stabilization of PduCDE under thermal stress is not an artefact of molecular crowding but a result of the interaction between the two proteins.[89] In the presence of PduBB' a ~5°C shift in the melting temperature (T_m) of PduCDE is observed **(Fig. 3.8a)**. A similar shift in melting point is not observed in the presence of PduBM38L or PduB'.
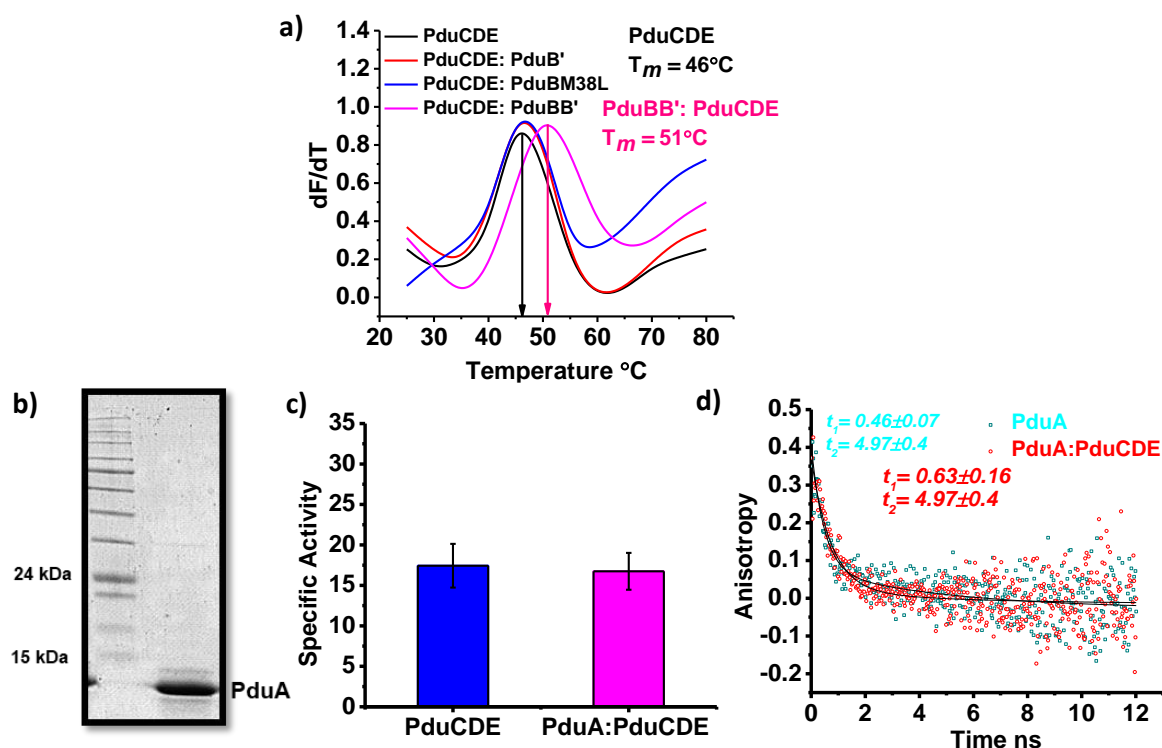
Fig. 3.8: a) Thermal shift assay of PduCDE in the absence and presence of the three forms of PduB shell proteins mixed in shell protein: enzyme ratio of 8:1 molar ratio. For thermal shift assay, 8-Anilinonaphthalene-1-sulfonic acid (*ANS*) dye was used to probe

the unfolding of PduCDE in the absence and presence of the shell proteins. The plot shows 1st derivative of fluorescence intensity at 488 nm as a function of temperature, b) SDS PAGE of PduA, c) Specific activity of PduCDE post thermal stress at 45°C in the absence and presence of PduA at shell protein: enzyme molar ratio of 64:1, d) Fluorescence anisotropy of PduA in the absence and presence of PduCDE mixed in equimolar ratio.

3.7 Conclusion

The present work elucidates the link between molecular confinement and conformational stability of enzymes in PduMCP. We provide evidence in support of the idea that the major shell protein PduBB' has the ability to improve the catalytic activity and chaperone the enzyme PduCDE under thermal stress. The protective role of PduBB' is reinforced by the fact that the activity of PduMCP is less affected by thermal perturbations compared to the bare enzyme. This strengthens the idea that encapsulation indeed influences the catalytic efficiency of enzymes within PduMCP [90]

Through our work we show that the simultaneous expression and assembly of the PduB and PduB' is very important for the formation of a stable and a soluble shell protein and the combination has a stronger affinity towards PduCDE compared to the components PduB and PduB'. The N-terminal region of PduBB' has a crucial role is providing solubility to the shell protein. It is likely that the extended N-terminal region in PduBB' or PduBM38L causes a steric hindrance, preventing their self-association and improving their solubility. However, it is important to note that mere presence of N-terminal region in shell protein is not sufficient to bind and protect PduCDE. PduBM38L, despite having the extended N-terminal region fails to influence the catalytic activity of the enzyme. This may be because this shell protein is not structurally stable and has a low tendency to self-assemble, limiting its ability to scaffold PduCDE. In a nutshell, the absence of N-terminal region in PduB' renders it insoluble, preventing it from firmly binding and protecting PduCDE. PduBM38L, on the other hand, despite having an N-terminal region and being soluble, lacks structural stability and has the least self-assembly property. We therefore, conclude that the presence of an N-terminal region, thermo-resistance, and naturally optimized self-assembly behavior give PduBB' its chaperone-like activity toward PduCDE.

It should be noted that in a confined PduMCP environment, a specific enzyme may have multiple shell protein partners. However, *in vitro*, the affinity of shell proteins for enzymes and their impact on their stability may differ. For example, under the conditions described for

PduBB', another major shell protein, PduA, does not interact with or affect PduCDE (**Fig. 3.8, c-e**). In the context of future application, the thermal stability of PduBB' is indeed encouraging and may prove to be useful in the development of shell protein based 2-D and 3-D thermostable nano-materials using genetic engineering and chemical biology approaches. These all-protein smart biomaterials can be an alternative to the confinement of enzymes via immobilization onto artificial matrices like silica [91], glyoxyl-agaraose [92] or epoxy-activated supports [93].

Additional note 3.1:

Where did mesophilic bacteria like Salmonella acquire such a heat-resistant shell protein?

*The high thermal stability of the shell protein PduBB' and its chaperone like activity intrigues us to look for the evolutionary origin of such a thermostable shell protein in a mesophilic Salmonella enterica. The fact that Salmonella enterica is a mesophilic bacteria led us to question how come it ended up acquiring such thermostable protein. We came across an article where authors discovered genes for PduMCP are found in thermophilic Parageobacillus species. Motivatd by this finding, we went for our own bioinformatics analysis and a BLAST search using the amino acid sequence for PduB from Parageobacillus thermoglucosidacius produced significant alignments with PduB from 9 thermophilic bacterial species (**Box 3.2a**) They belong to the families Bacillaceae, Sporomusaceae, Thermoanaerobacterales Family III. Incertae Sedis, Peptococcaceae and Clostridaceae that fall under phylum Firmicutes. Salmonella on the other hand belongs to the family Enterobacteriaceae which falls under the phylum Proteobacteria. Divergence tree showed that the ancestors of Salmonella diverged from those of the thermophiles around 3000 million years ago (**Box 3.2b**). Yet, pair wise sequence alignment using BLAST reveals more than 60 % identity between PduBB' from Salmonella PduB from thermophilic Firmicutes (**Box 3.2a**). This suggests that the thermostable property of the shell protein PduBB' has been passed down to Salmonella enterica by its thermophilic ancestors. As a matter of fact, we find an excellent comparative genomic analysis by Ravcheev et al. 2019 in which the authors show that PduMCP is not specific to any single phylum. It first appeared among the Firmicutes and was horizontally transferred to other groups [94]. It appears that the thermostable property of shell protein might have helped in conserving the activity of Pdu enzymes in thermophilic Firmicutes under heat stress.*

Box 3.2

a)

Species	Family	% identity of PduB
Parageobacillus thermoglucosidasius	Bacillaceae	67.78%
Sporolituus thermophilus	Sporomusaceae	65.17%
Thermoanaerobacterium thermosaccharolyticum	Thermoanaerobacterales Family III. Incertae Sedis	61.25%
Desulfofundulus thermocisternus	Peptococcaceae	63.81%
Thermincola ferriacetica	Peptococcaceae	61.19%
Thermincola potens	Peptococcaceae	61.57%
Desulfofundulus thermosubterraneus	Peptococcaceae	71.56%
Quasibacillus thermotolerans	Bacillaceae	62.78%
Thermohalobacter berrensis	Clostridaceae	65.04%

b)
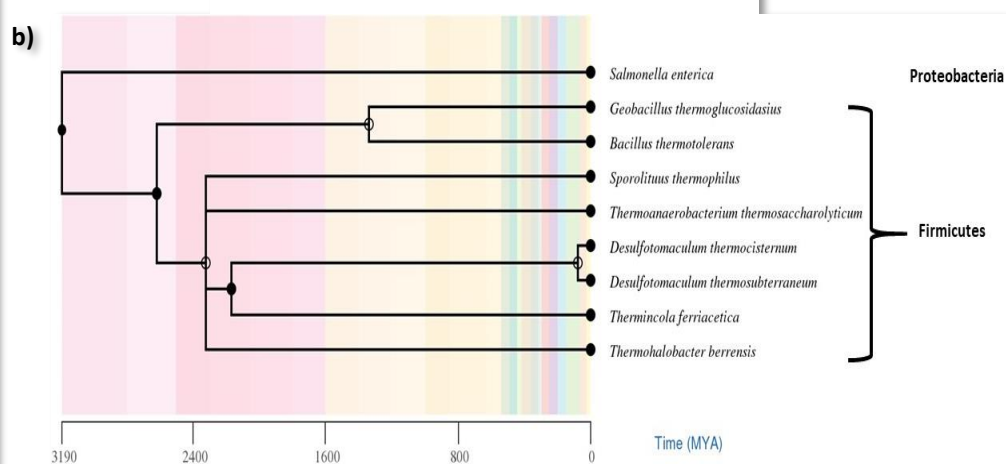
a) List of thermophilic species with *PduB* gene, their families and percentage identity between PduB of thermophiles with that of *Salmonella enterica*, b) Time line for divergence of Salmonella enterica and other thermophilic firmicutes.

Note:

Chapter 4

Factors Governing the Self-Assembly of Shell Protein and Enzyme

4.1 Introduction

In the previous chapter, I investigated the impact of compartmentalization on the enzymatic function of PduMCP. The experimental results showed that the outer shell of PduMCP has the potential to protect the encapsulated enzyme from denaturation under thermal stress. The major shell protein PduBB'balances the self-assembly properties of its component proteins (PduB and PduB') and provides interaction based thermal stability to the native enzyme PduCDE. The results helped me to assign a chaperone like character to shell protein PduBB. While working on this concept, I got fascinated by the fact that the shell protein PduBB' is a combination of not only two shell proteins but also each one's distinct self-assembly properties. This prompted me to use PduBB'to address my next set of objectives which were to identify the factors that govern the self-assembly of PduMCP proteins in solution phase. Available reports in literature suggest that the self-assembly of PduMCP is the result of several genetic, biochemical, and physical stimuli orchestrating inside the bacterial cell [46] [47, 95] [96]. However, it is important to understand how multiple subunits of proteins are driven towards one another and orchestrated into small yet complex catalytic reactors.

In my attempt to find a solution to this problem, I came across the concept of phase separation, which has been proposed as the underlying mechanism driving the self-assembly and formation of a number of membraneless organelles. It is remarkable to think about proteins separating out of solution and behaving like liquids. Protein phase separation is a dynamic process that is influenced by the surrounding environmental conditions such as molecular crowding or ionic strength. Whether or not phase separation played any role in the self-assembly of PduMCP is indeed an idea worth exploring. Interestingly, initial bioinformatics analysis of the shell proteins of PduMCP using catGranule server [97] suggested that out of the 8 shell proteins of PduMCP, the major shell proteins PduB and PduB' have a high phase separation propensity (**Fig. 4.1a**). Although the crystal structure of PduBB' from Salmonella enterica is not solved, the available crystal structure of a homologous protein, PduB from *Lactobacillus reuteri* (**Fig. 4.1b**) [39], suggests that PduB is a trimeric shell protein and ionic interaction are involved in the lateral associations of the trimeric shell protein assemblies. It would be interesting to study how ionic strength of the surrounding environment as well as molecular crowding impact the self-assembly behavior of PduBB'. In this chapter we present our results on the self-assembly of shell protein PduBB' in a crowded environment and under different salt concentrations. We also probe the co-assembly of PduBB' with PduCDE and its consequence on the catalytic efficiency of the

enzyme. We find that a major shell protein PduBB' tend to self-assemble under macromolecular crowded environment and suitable ionic strength. Microscopic visualization and biophysical studies reveal phase separation to be the principle mechanism behind the self-association of shell protein in the presence of salts and macromolecular crowding. The shell protein PduBB' interacts with the enzyme diol-dehydratase PduCDE and co-assemble into phase separated liquid droplets. The co-assembly of PduCDE and PduBB' results in the enhancement of catalytic activity of the enzyme. The shell proteins that make up PduBB' (PduB and PduB') have contrasting self-assembly behavior. While N-terminal truncated PduB' has a high self-associating property and forms solid assemblies that separates out of solution, the longer component of the shell protein PduBM38L is more soluble and show least tendency to undergo phase separation. A combination of spectroscopic, imaging and biochemical techniques shows the relevance of divalent cation Mg^{2+} in providing stability to intact PduMCP *in vivo*. Together our results suggest a combination of protein-protein interactions and phase separation guiding the self-assembly of Pdu shell protein and enzyme in solution phase.

4.2 Phase separation of shell protein PduBB' in a crowded environment

The cellular microenvironment is highly viscous and crowded due to the presence of many bio-macromolecules[98-100]. It is therefore important that we study the self-assembly dynamics of proteins under crowded environment. Polyethylene glycol (PEG) is a well-known crowding agent which has been extensively used for mimicking the cellular microenvironment *in vitro*.[101, 102] To visualize the dynamics of PduBB' under crowded environment, fluorescence microscopy of Alexa-488 labelled PduBB' solution is performed in the presence PEG-6000 (5% w/v) and Mg^{2+}, Na^+, K^+ salts. These salts are selected as they are essential for production and purification of PduMCP from *Salmonella enterica* LT2[55]. Under fluorescence microscope, distinct PduBB' assemblies are seen in the presence Mg^{2+} salts (**Fig. 4.1c, i**). Careful observation of the dynamics of the PduBB' assemblies in real time, reveal their liquid like behavior, undergoing fusion in solution phase (**Fig. 4.1c, ii**). We observe that $MgSO_4$ and $MgCl_2$ are able to trigger the LLPS of PduBB' in a range of concentrations (**Fig. 4.1c, i**). In comparison to Mg^{2+} salts, monovalent salt, namely NaCl induce LLPS of PduBB' at high concentrations of 200 mM and above (**Fig. 4.1c, iii),** while KCl is able to induce LLPS in PduBB' only at higher concentration of 400 mM (**Fig. 4.1c, iii)**. The potential of the salts to trigger LLPS of PduBB' follows the following order: $MgSO_4$ >$MgCl_2$ >NaCl > KCl.
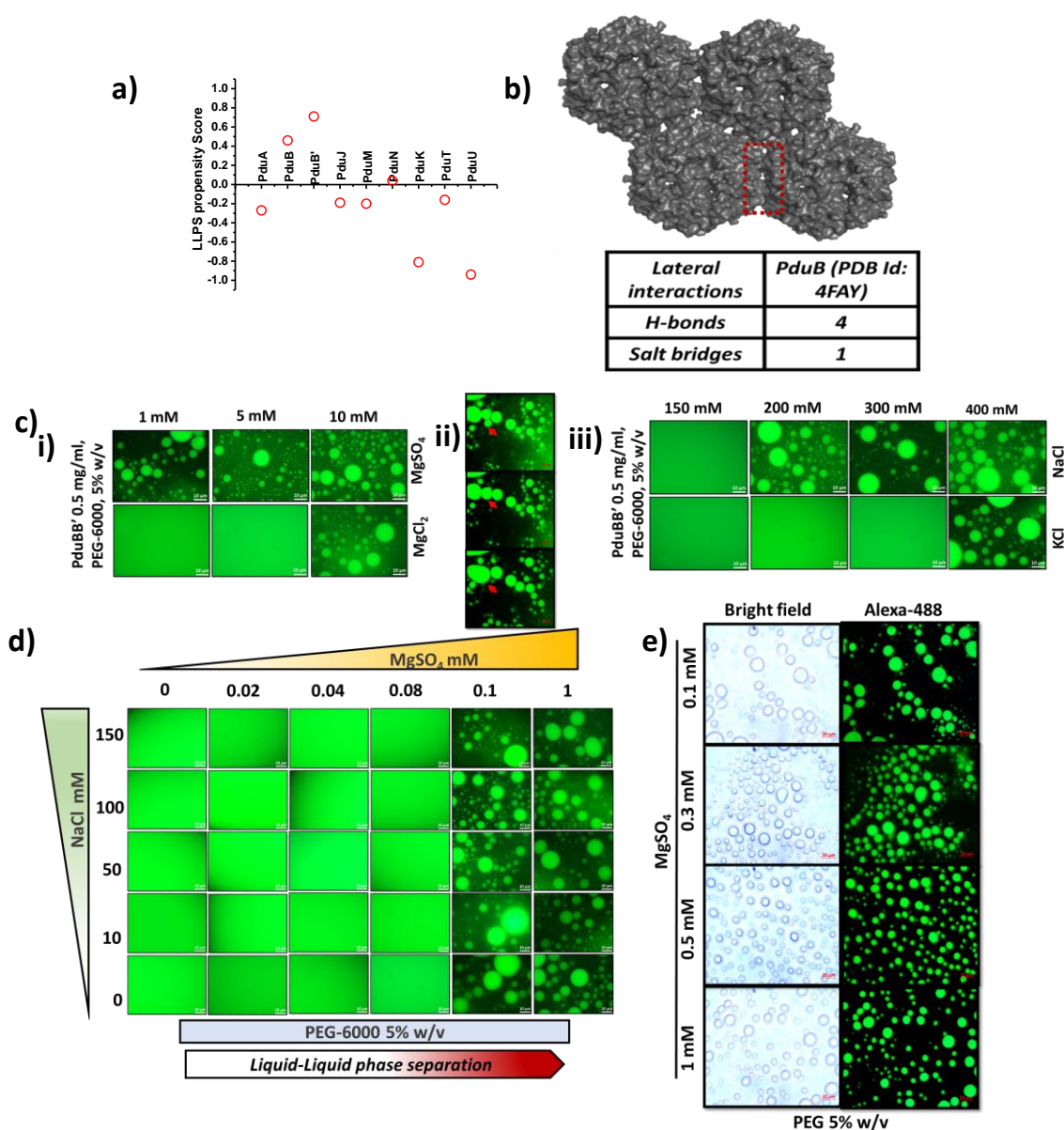
Figure 4.1: a) Liquid-liquid phase separation propensity of Pdu shell proteins predicted using catGranule, b) Edge-Edge lateral association in shell protein PduB from *Lactobacillus reuteri* **(PDB ID: 4FAY), c) Microscopic visualization of self-assembly dynamics of PduBB' in solution phase in the presence of salts and crowding agent (PEG-8000, 5% w/v): (i) Liquid-liquid phase separation of PduBB' in the presence of MgSO$_4$ (at 1 mM, 5 mM and 10 mM concentrations), and MgCl$_2$ (at 10 mM concentration), (ii) Fusion of protein droplets confirming liquid nature of the PduBB' droplets, (iii) Liquid-liquid phase separation of PduBB' in the presence of NaCl (at 200 mM, 300 mM and 400 mM concentrations) and KCl at 400 mM concentration, d) Phase diagram showing phase separation of PduBB' (0.5 mg/ml) in solution phase at different concentrations of NaCl and MgSO$_4$, e) Liquid-liquid phase separation of PduBB' at MgSO$_4$ concentrations between 0.1 mM to 1mM**

Further we see that, Mg^{2+} alone is sufficient to trigger the LLPS of shell protein PduBB' and the presence of NaCl is not necessary for the same **(Fig. 4.1d-e).** The LLPS of PduBB' protein with increase in solution ionic strength suggests a role of hydrophobic interactions in bringing the shell protein molecules together. As shell proteins have exposed hydrophobic patches on its surface[61], the presence of kosmotropic ions like Mg^{2+} would expel out water molecules from the surrounding environment of PduBB', bringing the shell protein molecules closer due to hydrophobic interactions. Since Na^+ ions are weak kosmotrope as compared to Mg^{2+}, a high concentration of NaCl is required to trigger LLPS of PduBB'. Dye alone does not show any sign of droplet formation under similar condition **(Fig. 4.2a, i)** and clear transparent droplets are observed in case of unlabeled PduBB' sample **(Fig. 4.2a, ii)**, indicating that fluorophore labeling does not play any role in LLPS of shell protein unlabeled PduBB'. We also observe that only PEG is unable to induce LLPS of PduBB', indicating that salts are essential for the formation of shell protein liquid droplet **(Fig. 4.2b)**. These liquid phases are found to be stable at 37°C for several days without any aggregation **(Fig. 4.2c)**.
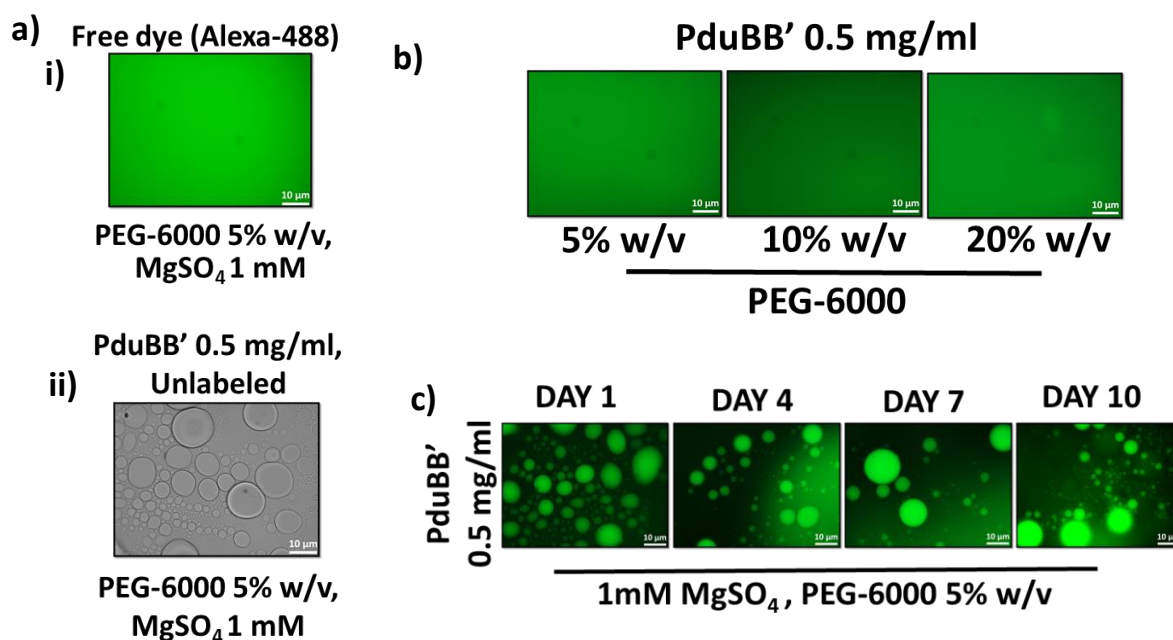
Figure 4.2: a) Microscopic visualization of (i) Free dye alexa-488 and (ii) unlabeled PduBB' in the presence of crowding agent PEG-8000 5% w/v and 1 mM MgSO₄, b) Microscopic visualization of Alexa-488 labeled PduBB' in in the presence of 5% w/v-20% w/v PEG-6000, c) Day dependent microscopic visualization of PduBB' droplets after incubating the protein samples at 37°C.

Next, we check for the self-assembly behavior of PduBB' in the presence of salts withour adding PEG. In the absence of the crowding agent PEG, salts alone do not trigger LLPS but tend to increase the size of PduBB' assemblies after 1 hr or incubation.

To understand how change in ionic strength of environment affects the self-assembly behavior PduBB', we look at the size distribution of PduBB' (0.1 mg/ml) in solution phase without adding any crowding agent, in the absence and presence of NaCl (10 mM-150 mM) using dynamic light scattering (DLS) **(Fig. 4.3a)**. In solution, PduBB' (0.1 mg/ml) shows a wide size distribution of shell assemblies (intensity percentage in DLS) between 10nm-500nm. This suggests that a heterogeneous population of shell protein assemblies exists in solution phase. However, after 1 h of incubation in the presence of NaCl, NaCl self-association of PduBB' occurs in solution. At 50 mM NaCl, more than 90% of the size distribution is seen above 500 nm. Increase in NaCl concentration to 150 mM, results in further self-association of the shell protein with a majority of PduBB' assemblies falling in the micron range. A similar trend is also noticed when PduBB' is incubated in the presence of KCl **(Fig. 4.3b).** The self-association behavior of PduBB' in the presence of NaCl or KCl becomes more apparent when we look at the maximum population (number percentage in DLS) of the shell protein assemblies in solution at different ionic strengths **(Fig. 4.3c).** In the absence of NaCl or KCl, the solution containing PduBB' shows a maximum population of shell protein assemblies around 10 nm which changes to a maximum population in the micron range upon increasing the salt concentration to 150 mM.

Interestingly, the effect of Mg^{2+} salts on PduBB' assembly is found to be more aggressive than the monovalent salts (NaCl and KCl), with PduBB' showing maximum population of size above 1000 nm at 1 mM and above 2000 nm at 10 mM of $MgCl_2$ or $MgSO_4$ **(Fig. 4.3d).** Our data indicates that divalent salts ($MgCl_2$ and $MgSO_4$) have a higher potential than monovalent salts (NaCl and KCl) to promote self-association of PduBB' in solution. This observation is similar to the one seen in the presence of crowding agent PEG, and hints towards the kosmotropic effect of salts in triggering the self-assembly of shell protein PduBB'.
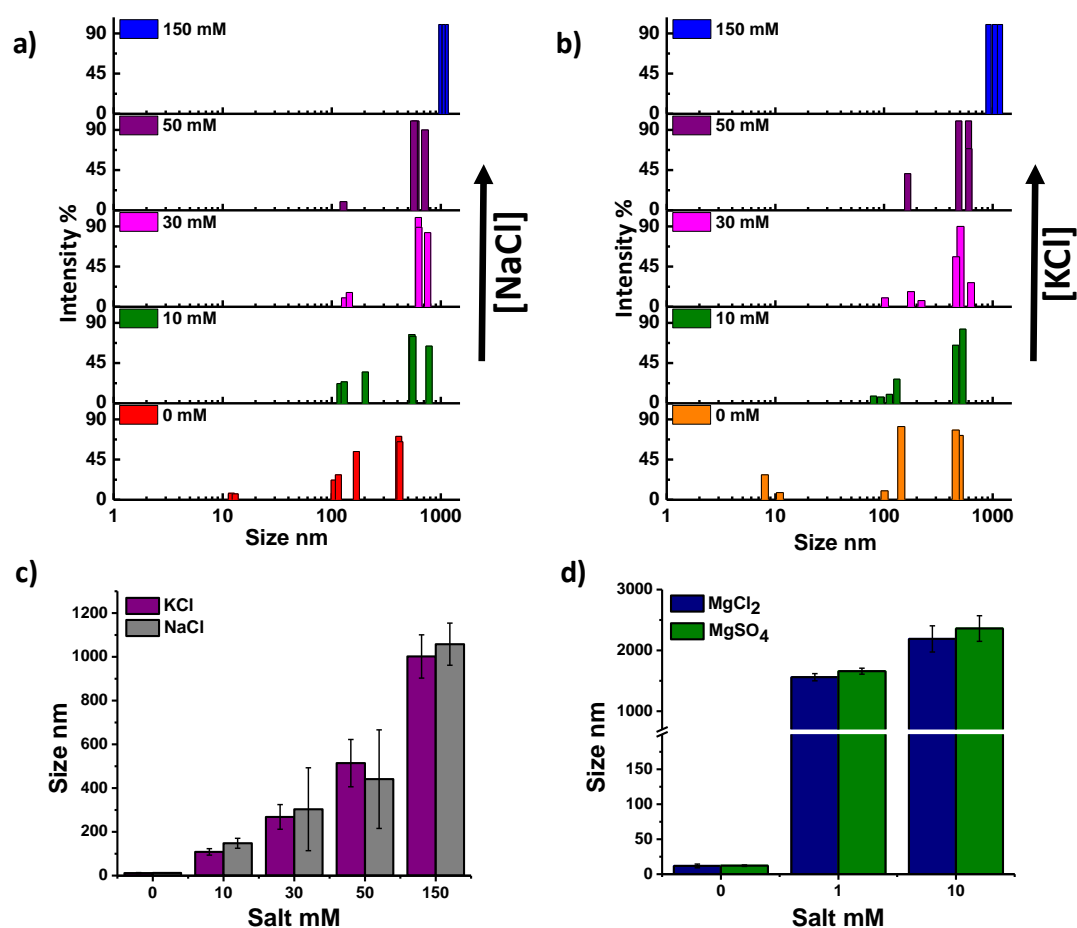
Figure 4.3: a) Size distribution of shell protein PduBB' under NaCl concentrations varied from 0 mM to 150 mM, b) Size distribution of shell protein PduBB' under KCl concentrations varied from 0 mM to 150 mM, c) Size of PduBB' assemblies in solution phase in the presence of KCl and NaCl at concentrations ranging from 0 mM to 150 mM, d) Size of PduBB' assemblies in solution phase in the absence and presence of MgCl$_2$ and MgSO$_4$ (1mM and 10 mM).

Purified PduBB' has an overall negative surface charge giving zeta potential value of -34.2 mV. Presence of salts increase the zeta potential value of PduBB' solution indicating surface charge masking of the shell protein by metal ions present in the vicinity of PduBB' **(Fig. 4.4a)**. This would reduce the electrostatic repulsion among the shell proteins inducing their self-association **(Fig. 4.4a)**. Stronger surface charge masking of PduBB' by Mg^{2+} ions explains why Mg^{2+} salts trigger the self-association of PduBB' at very low concentration of 1 mM. This suggests that beside kosmotropic effect of salts, surface charge masking of PduBB' molecules in the presence of metal ions, also triggers their self-assembly.

The resultant PduBB' assemblies are seen as solid associates under the microscope. To check if these assemblies are soluble associates or irreversible aggregates, the protein samples are dialyzed in 10 mM phosphate buffer (pH 7.4) to remove the salts. Post dialysis, the PduBB' solution becomes clear and we do not see any assembly under microscope **(Fig. 4.4b).** Alexa-488 labelled PduBB', in the presence of salts showed no substantial difference in fluorescence lifetime indicating retention of structure upon salt induced self-assembly **(Fig. 4.4c).** Circular dichroism studies of salt treated PduBB' samples post dialysis in 10 mM phosphate buffer (pH 7.4) also show the preservation of the secondary structure **(Fig. 4.4d),** confirming the benign effect of salts on the secondary structure of the protein.
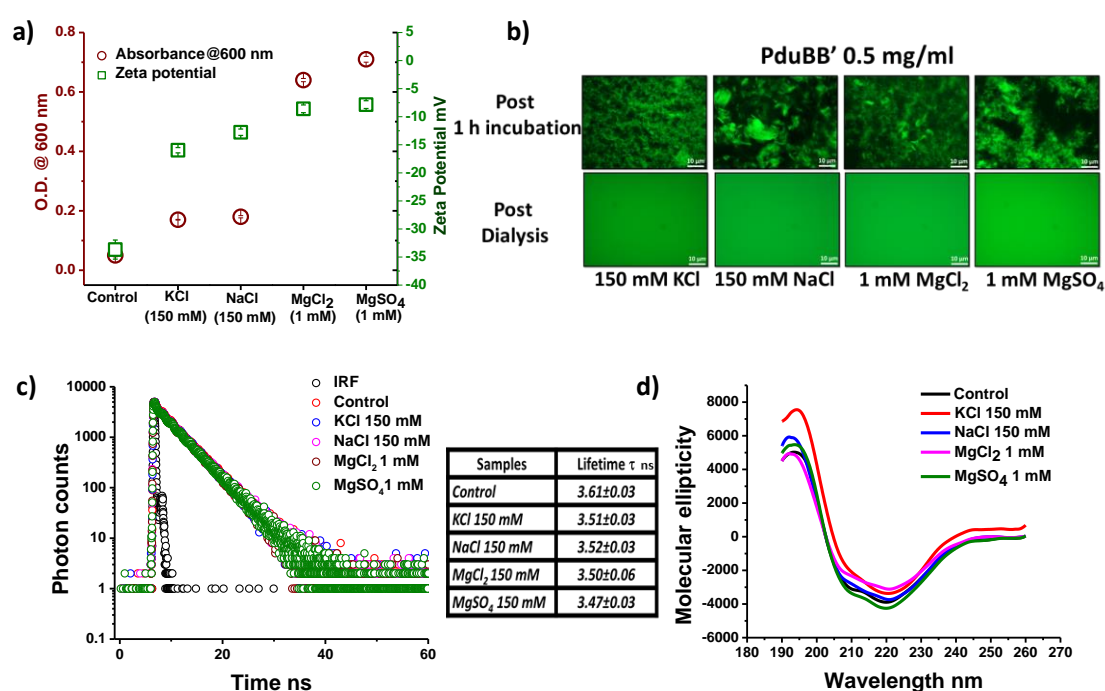
Figure 4.4: a) Absorbance at 600 nm and zeta potential of PduBB' solution in the absence and presence of salts, b) Microscopic visualization of assemblies of Alexa-488 labeled PduBB' (0.5 mg/ml) in solution phase in the presence of salts. PduBB' tend to self-associate within an hour forming visible soluble associates which disappear upon dialysis and removal of salts, c) Fluorescence lifetime of Alexa-488 labeled PduBB' (0.5 mg/ml) in the absence and presence of salts, d) CD spectroscopy of PduBB' (0.5 mg/ml) in the absence and presence of salts post dialysis.

Further, a combination of lifetime measurement studies, and fluorescence anisotropy experiments show that during the process of phase separation, the backbone of the PduBB' protein remains solvated while the dynamics of the background are restricted **(Fig. 4.5)**. Solvation of protein backbone would maintain the protein molecules in liquid phase and

prevent aggregation. The restricted backbone flexibility may occur due to the increased solution viscosity as a result of macromolecular crowding.

We determine fluorescence lifetime of Alexa-488 labelled PduBB' under control and phase separated condition (1 mM MgSO4, 5% w/v PEG-6000) **(Fig. 4.5a, i)**. Fitting the first 10 ns seconds of lifetime decay curve using monoexponential decay equation, shows a small decrease in lifetime post phase separation by ~0.5 ns **(Fig. 4.5a, ii).** This change in lifetime may be due to the alteration in the microenvironment of fluorophore molecule due to local enrichment of ions which tend to mask the surface charge of the shell protein during phase separation. We next perform solvent accessibility of Alexa-488 labelled PduBB' under control and phase separated condition, using water soluble quencher potassium iodide (KI) **(Fig. 4.5b, i-ii).** In both cases the fluorescence lifetime of labelled PduBB' decreases with the increase in KI concentration. A plot of τ_0/τ at different concentrations of KI (Stern-Volmer plot) where τ_0 is the fluorescence lifetime of Alexa-488 labeled PduBB' at 0 mM KI and τ is the fluorescence lifetime of Alexa-488 labeled PduBB' at different concentrations of KI, is linear in both solution and phase separated condition **(Fig. 4.5c)**. Fitting the Stern-Volmer plots using linear fit equation shows same binding constants for the quencher in case of both control and phase separated PduBB' sample. This suggests that the accessibility of KI to the shell protein backbone is similar in case of both control and phase separated samples. In other words, phase separation does not result in disruption of the solvation layer of the shell protein backbone. The solvation of peptide backbone of PduBB' may be the reason why the shell protein behaves like liquid post phase separation and not collapse into aggregates. Further, in time resolved anisotropy studies we observed a rotational correlation time of 1.7±0.1 ns in solution, and of 5.2±0.5 ns in LLPS PduBB' **(Fig. 4.5d)**. The higher rotational correlation time, suggests a restricted rotational dynamic of the shell protein in the crowded micro-environment. Together our fluorescence lifetime and anisotropy studies indicate that during the LLPS process the shell protein backbone remains solvated but have restricted rotational dynamics.
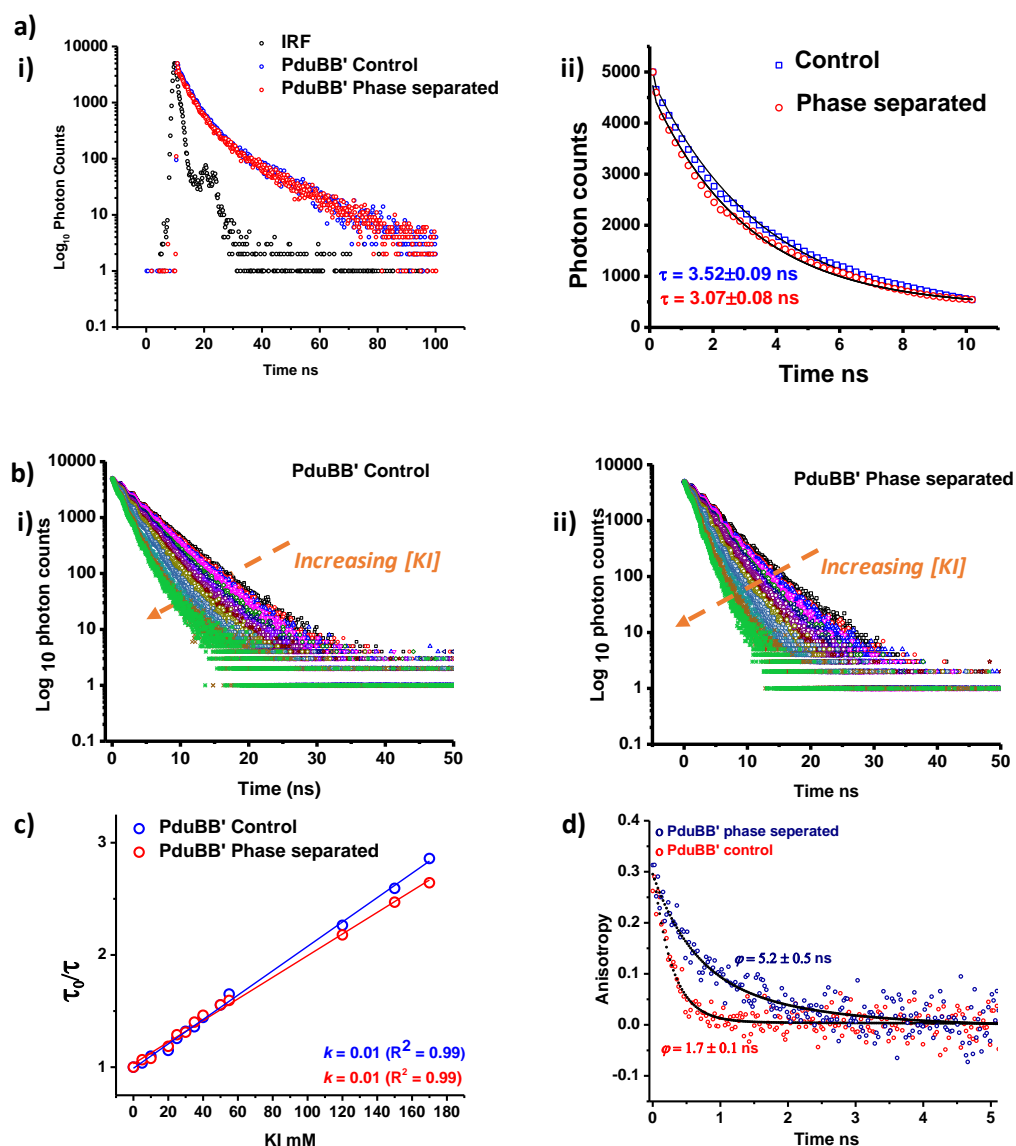
Figure 4.5: a) Fluorescence lifetime of Alexa-488 labeled PduBB' (0.5 mg/ml) under (i) control and (ii) phase separated condition (1mM MgSO₄ and 5% w/v PEG-6000), b) Fluorescence lifetime of PduBB' under (i) control and (ii) phase separated condition (1mM MgSO₄ and 5% w/v PEG-6000) in the presence of increasing concentrations of potassium iodide (KI), c) Stern-Volmer plot of τ_0/τ (τ_0 = fluorescence lifetime of Alexa-488 labeled PduBB' at 0 mM KI, τ = fluorescence lifetime of Alexa-488 labeled PduBB' at different concentrations of KI, d) Time resolved fluorescence anisotropy of PduBB' under control environment and post phase separation.

4.3 PduBB' has higher tendency than PduCDE to undergo phase separation

Next we compare how phase separation of PduBB' is different from a native enzyme of PduMCP. For this study we selected diol dehydratase PduCDE which is the signature enzyme

of PduMCP. For comparing the LLPS behavior PduCDE and PduBB', we look for the minimum concentrations of proteins to observe LLPS in solution phase. The LLPS of PduBB' is observed up to very low shell protein concentration of 0.025 mg/ml while the enzyme undergoes phase separation only at a concentration of 1 mg/ml **(Fig. 4.6a)** suggesting that PduBB' has a higher phase separation propensity than the enzyme PduCDE. The same is predicted by catGranule server that shows high LLPS propensity of shell protein as compared to PduCDE **(Fig. 4.6b).** Compared to the enzyme PduCDE, which is globular in nature, shell protein PduBB' has more exposed hydrophobic patches on its surface[61]. The hydrophobic dye ANS shows higher fluorescence intensity upon interacting with native PduBB' compared to PduCDE at same protein concentration suggesting more exposed hydrophobic patches for PduBB' compared to PduCDE **(Fig. 4.6c)**. Salts like $MgSO_4$ are kosmotropic in nature and tend to retain water molecules. Because PduBB' has more exposed hydrophobic patches than the enzyme, the kosmotropic effect of salt would be greater in the case of the shell protein, driving PduBB' LLPS at much lower concentration as compared to enzyme. Bio-layer interferometry studies also show that PduBB':PduBB' interaction is stronger than PduCDE:PduCDE interaction **(Fig. 4.6, d-e).** For the PduBB': PduBB' interaction, no dissociation is observed post association event, indicating very strong association of the molecules **(Fig. 4.6d)**. A global fitting of the association and dissociation kinetics of PduCDE:PduCDE interaction gives a dissociation constant (K_d) of 3.5 x 10^{-6} M **(Fig. 4.6e).** Strong affinity of shell protein molecules towards one another may be a contributing factor to their higher tendency to undergo self-association and phase separation.

In the presence of macromolecular crowding (5% PEG-6000) and 1mM MgSO4, PduBB' and PduCDE, co-localize into protein droplets **(Fig. 4.7a)**. PduCDE co-phase separated with PduBB' shows enhanced catalytic rate compared to PduCDE alone **(Fig. 4.7b)**. This result highlights the importance of both phase separation and shell protein: enzyme interaction in improving the catalytic activity of the enzyme PduCDE. We also observe that PduCDE could be recruited within pre-formed phase separated PduBB' droplets **(Fig. 4.7c).** Although PduBB' displays a strong LLPS propensity at wide concentration range, it undergoes liquid-solid transition forming solid assemblies at high protein concentration of 2 mg/ml. These solid shell protein assemblies do not show fluidic nature like protein droplets but retain the ability to recruit PduCDE molecules. A control study using BSA as a non-native protein negates possibility of non-specific interaction in our study **(Fig. Box 4.1).**
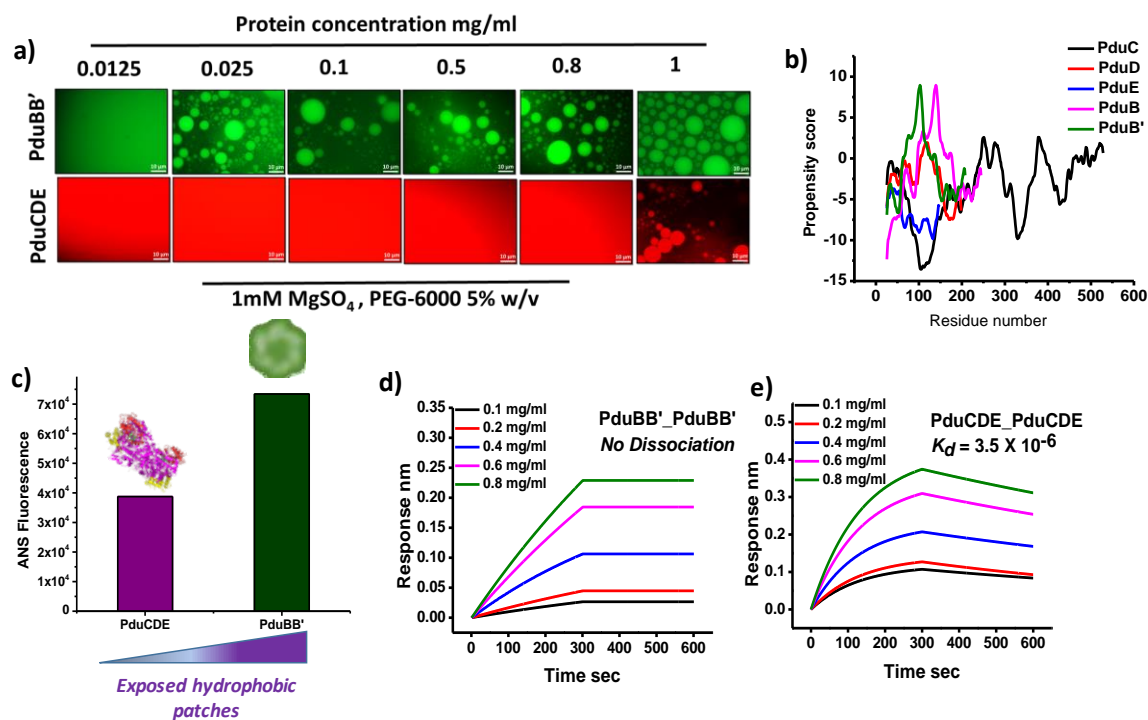
Figure 4.6: a) Liquid-liquid phase separation of PduBB' and PduCDE at protein concentrations 0.0125-1 mg/ml, b) LLPS propensity scores of the three subunits of PduCDE and the monomeric units of shell protein PduBB', c) ANS fluorescence of PduCDE and PduBB' at 0.5 mg/ml protein concentration, d) Protein-protein interaction performed using biolayer interferometry to show association and dissociation kinetics for PduBB': PduBB' association, e) Protein-protein interaction performed using biolayer interferometry to show association and dissociation kinetics for PduCDE: PduCDE interaction.

In soluble form without any macromolecular crowding and salt, the enzyme PduCDE shows a specific activity of 23.3 ± 1.4 µmol min^{-1} mg^{-1}. We look for the specific activity of PduCDE at different concentrations (0.5-2 mg/ml) under phase separated condition, in the absence and presence of PduBB' **(Fig. 3f)**. In the absence of PduBB', the enzyme PduCDE shows an activity of ~ 25.4 ± 0.05 µmol min^{-1} mg^{-1} at protein concentration of 0.5 mg/ml **(Fig. 3f)**. An increment in specific activity of PduCDE is seen only at enzyme concentrations 1 mg/ml and above (32.2 ± 0.5, 32.5 ± 0.5 and 37.7 ± 0.1 µmol min^{-1} mg^{-1} at 1, 1.5 and 2 mg/ml concentrations respectively). Since PduCDE fails to undergo phase separation below 1 mg/ml, at 0.5 mg/ml it exhibits an activity similar to the one seen in the absence of salt and crowding agent. However, in the presence of shell protein PduBB', PduCDE shows higher specific activity as compared to when it is phase separated alone, indicating the significance

of PduBB' in improving the catalytic function of the signature enzyme **(Fig. 4.7d).** Also, the specific activity is enhanced by ~3 fold when the enzyme is recruited within solid phase at 2 mg/ml concentration of PduBB' **(Fig. 4.7d).**
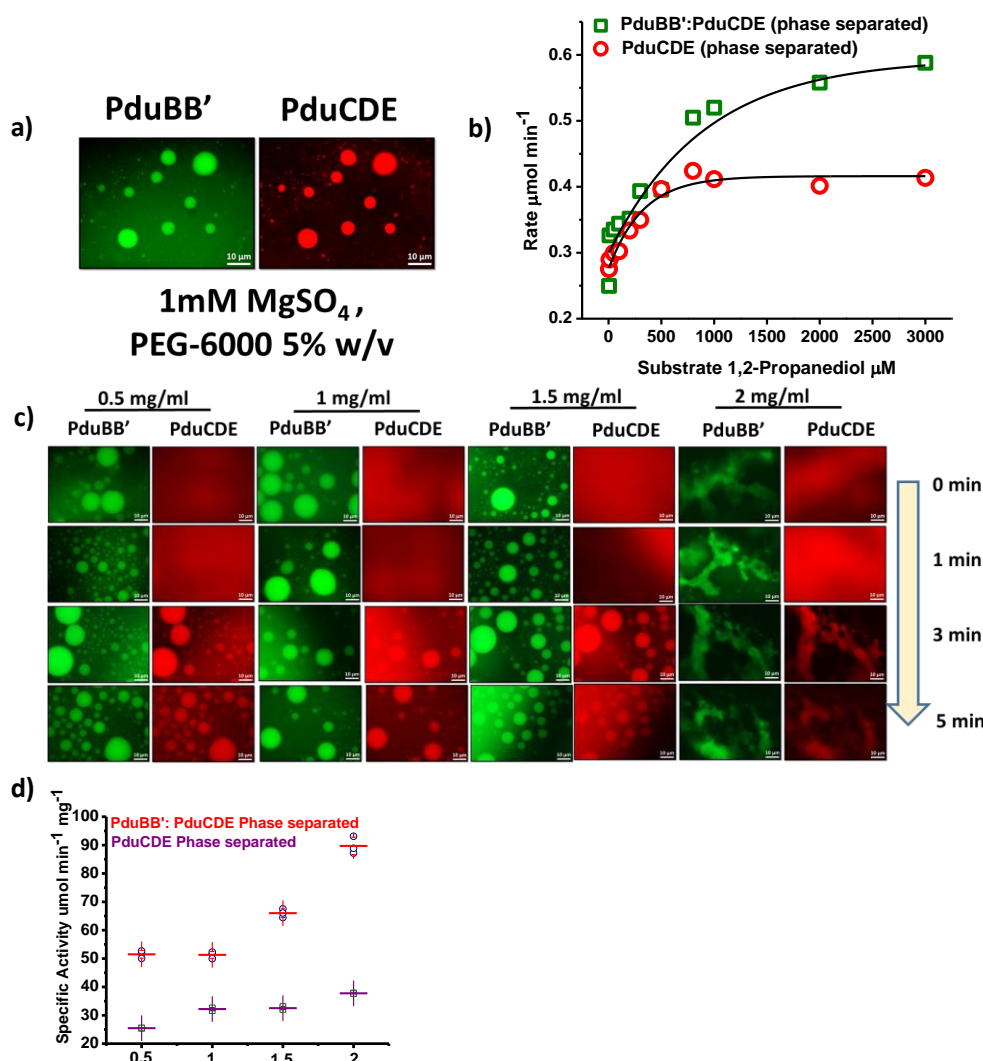
Figure 4.7: a) Co-localization of PduCDE (0.5 mg/ml) and PduBB' (0.5 mg/ml), b) rate of reaction at different substrate concentration for PduCDE and PduBB':PduCDE under phase separated condition, c) Time dependent recruitment of Texas red-labelled PduCDE within Alexa-488 labelled phase separated PduBB', d) Specific activity of PduCDE and PduCDE: PduBB' at different protein concentrations in the presence of 1 mM MgSO$_4$ and 5%w/v PEG-6000, Each data point shows three specific activity values (circles in case of PduBB': PduCDE and squares in case of PduCDE alone), along with mean (horizontal line) and standard error (vertical line) [* $p < 0.05$].

4.4 Significance of PduB and PduB' in regulating self-assembly of PduBB'

In our previous report, we demonstrated that in solution phase, the three shell proteins (PduBB'/PduB/PduB') have distinct self-assembly behavior [103]. PduB' and PduB were found to have low and high aqueous solubility respectively, and their combination PduBB' displayed a balanced self-assembly behavior. Here we compare the phase separation behavior of PduBB' with its individual component shell proteins. On comparing the phase separation behavior of the three shell proteins we observe that PduB' forms solid assemblies at all concentrations tested **(Fig. 4.8a)**. The longer version of the shell protein PduB (generated as a single mutant, PduBM38L) fails to undergo LLPS at lower protein concentrations and we only notice a few PduBM38L droplets at concentrations of 1.5 mg/ml and higher **(Fig. 4.8a)**. High fluorescence intensity in the surrounding bulk solution, on the other hand, suggests that a significant portion of PduBM38L do not phase separate and remain in solution. These observations substantiate that the extended N-terminal region provides more solubility to the shell protein and regulates its self-assembly in solution[103]. The combination of the two shell proteins in PduBB' results in a well-balanced self-assembly behavior, producing PduBB 'liquid droplets up to a very low shell protein concentration of 0.025 mg/ml. Similar to the observation made in Fig. 2e, at high concentration of 2 mg/ml, PduBB' undergoes liquid-solid transition forming micron size solid assembl1ies. Visually, these solid assemblies resemble the solid assemblies of PduB' **(Fig. 4.8a).** These observations suggest that PduBB' integrates the self-assembly property of both PduB and PduB'. At lower PduBB' concentration its ability to self-associate (due to the presence of PduB') is balance by the longer component PduB. This balance between the two shell proteins is probably responsible for LLPS behavior of PduBB' at lower concentrations and starts to display PduB' character only when the shell protein concentration is increased to 2 mg/ml.

Box 4.1 Recruitment study using non-native protein

Recruitment studies to show co-phase separation of PduBB'-PduCDE, PduBB'-BSA and PduCDE-BSA; While PduCDE gets recruited into PduBB' droplets, non-native protein BSA fails get recruited witin PduBB' or PduCDE droplets.

This control study using BSA as a non-native protein negates possibility of non-specific protein-protein interactions in phase separation

Although PduB', PduBM38L and PduBB' are able to co-localize with PduCDE **(Fig. 4.8b)**, their effect on the catalytic activity of PduCDE is different **(Fig. 4.8b)**. The component proteins of PduBB' (PduBM38L and PduB') do not show any impact on the catalytic activity of PduCDE. On the other hand, co-phase separation of PduCDE with PduBB' improves its activity **(Fig. 4.8c)**, underscoring the significance of PduB and PduB' combination in improving the catalytic performance of the signature enzyme. It is interesting to note that all the three variants of shell protein interact with the enzyme PduCDE with micromolar affinity, but only the combination PduBB' is able to enhance the activity of the enzyme [103].
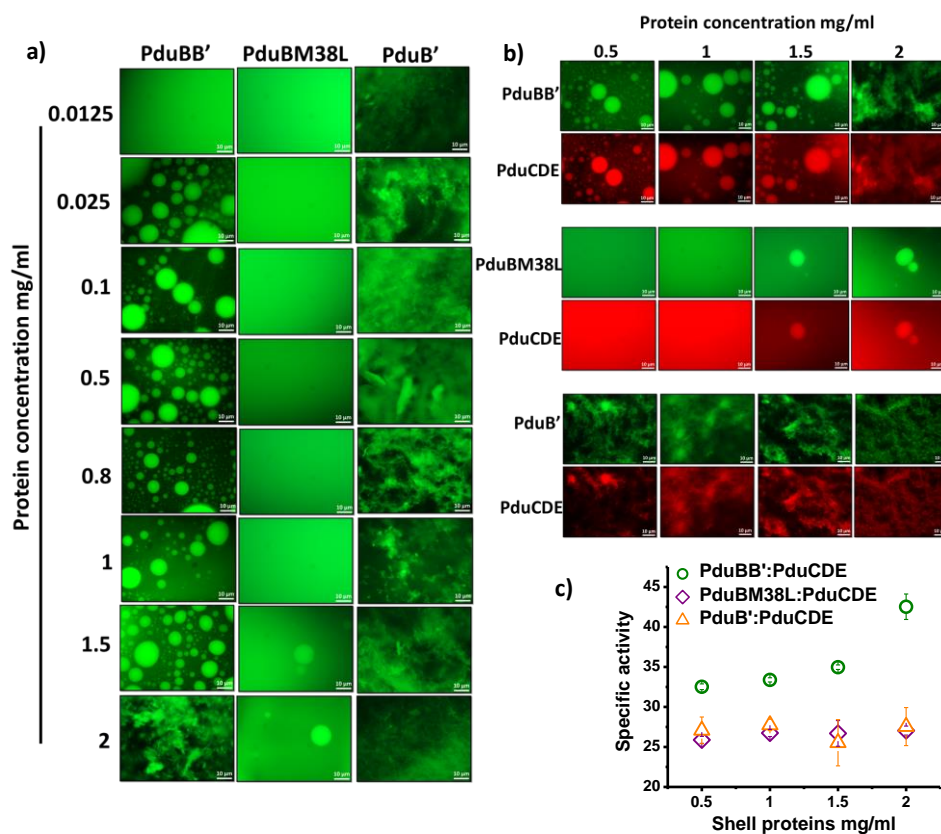
Figure 4.8: a) Phase separation behavior of PduBB', PduB' and PduBM38L at different protein concentrations, b) Co-localization of PduCDE with PduBB', PduB' and PduBM38L, c) Specific activity of PduCDE in the presence of PduBB', PduB' and PduBM38L under phase separated conditions.

4.5 Role of salts in the structure-function relationship of Pdu microcompartment

Since the presence of salts is observed to drive the self-association and phase separation of the main shell protein, we wonder whether the ionic strength of the environment has a substantial effect on the assembly, stability, and function of intact PduMCP. To address this, the diol dehydratase activity of purified PduMCPs is determined in the presence and absence of salts in its storage buffer (in 50 mM Tris buffer (pH 8.0) containing 50 mM KCl and 5 mM MgCl$_2$) **(Fig. 4.9a)**.

The salts are removed by dialyzing the PduMCPs against Tris buffer (pH 8.0) without salts. Post removal of salts by dialysis, the purified PduMCPs exhibit lower catalytic activity, **(Fig. 4.9a),** suggesting the importance of salts in maintain the optimum catalytic performance of purified PduMCP. Next we check if removal of Mg^{2+} alone affects the stability of intact PduMCP. For this study, purified PduMCPs are treated with a chelating agent, namely EDTA. The EDTA treated PduMCPs, are then tested for their thermal stability in the range of

20°C-90°C. **Fig. 4.9b** shows derivative plot of intrinsic fluorescence of PduMCP as a function of temperature, which gives us the melting temperature (T_m) of PduMCP. An increase in EDTA concentration results in the decrease in the T_m of PduMCP, suggesting lower stability of PduMCP in the presence of EDTA. The chelation of Mg^{2+} by EDTA also lowers the catalytic activity of the purified PduMCP. **(Fig. 4.9c).** It is however important to note that while 5 mM EDTA results in ~24% decrease in PduMCP activity, no further notable reduction in PduMCP activity is observed upon increasing the EDTA **(Fig. 4.9c)**. This suggests that Mg^{2+} chelation destabilizes PduMCP and reduces its activity, but it does not result in complete annihilation or denaturation of intact PduMCP. Size determination of the PduMCPs in the presence and absence of EDTA show no significant change in the hydrodynamic diameter of PduMCP, confirming that PduMCP remains intact upon chelation of Mg^{2+} by EDTA **(Fig. 4.9d)**.

Since Mg^{2+} salts have strong tendency to promote self-assembly of shell protein, it's possible that the Removal of Mg^{2+} would reduce the stability of the outer shell of PduMCP, lowering its thermal stability. However due to strong shell-shell interaction the outer shell may not totally disintegrate, maintaining the overall size of PduMCP in solution. Interestingly, we also observe that the chelation of Mg^{2+} by EDTA impairs the LLPS of PduBB': PduCDE mixture **(Fig. 4.9e),** suggesting that the absence of Mg^{2+} significantly affects the self-assembly behavior of Pdu proteins in solution. We also confirm that the presence of Mg^{2+} ions itself does not affect the catalytic activity of the enzyme PduCDE alone **(Fig. 4.9f),** suggesting that the Mg^{2+} ions only contribute towards the stability of PduMCP and do not influence the catalytic activity of the enzyme in any way.

Besides, Mg^{2+} salt is found to trigger the self-assembly and co-condensation of the major shell protein and enzyme in aqueous solution **(Fig. 4.10)**. The effect of 1 mM $MgSO_4$ on solutions of PduBB' and PduCDE is checked by turbidity assay. In the absence of the salt, both PduBB' and PduCDE (0.5 mg/ml each) gives clear solution without any turbidity **(Fig. 4.10a)**. In the presence of 1 mM $MgSO_4$ the solution of PduBB' turns turbid in an hour giving high optical density, while PduCDE solution remains clear **(Fig. 4.10a)**.
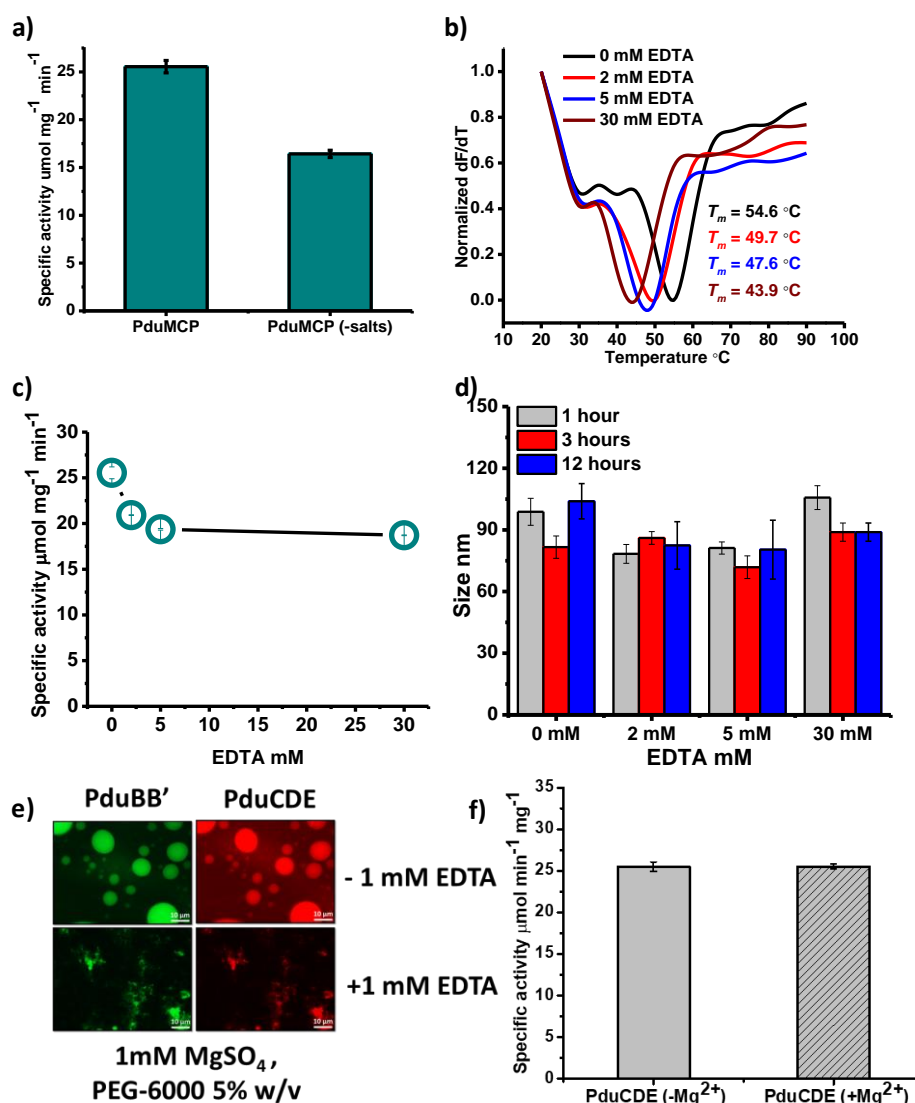
Figure 4.9: a) Specific activity of PduMCP in the presence and absence of salts (KCl and MgCl₂), b) Derivative plot of intrinsic fluorescence of PduMCP in the absence and presence of EDTA, c) Diol dehydratase assay of PduMCP in the absence and presence of EDTA, d) Hydrodynamic diameter of PduMCP in the absence and post treatment of EDTA for 1 h, 3 h and 12 h, c) Phase separation of PduBB' and PduCDE mixture in the absence and presence of 1 mM EDTA, scale bar = 10 µm, d) Specific activity of PduCDE in the absence and presence of 1 mM MgSO₄.

This suggests that shell protein PduBB' has a higher potential to undergo self-association in solution than the enzyme PduCDE. We check if the mixture of shell protein and enzyme has the ability to co-separate out in solution. For this study, we mix the shell protein PduBB' and enzyme PduCDE (0.5 mg/ml each in final mixture) in the presence of 1 mM MgSO₄, and condensate formed at different time interval is run on SDS-PAGE gel **(Fig. 4.10b)**. We

observe that, with time the intensity of the bands corresponding to PduBB' and PduCDE inccreases on gel. This indicates that in the presence of 1 mM MgSo4 triggers the self-assembly of PduCDE and PduBB' in solution phase.
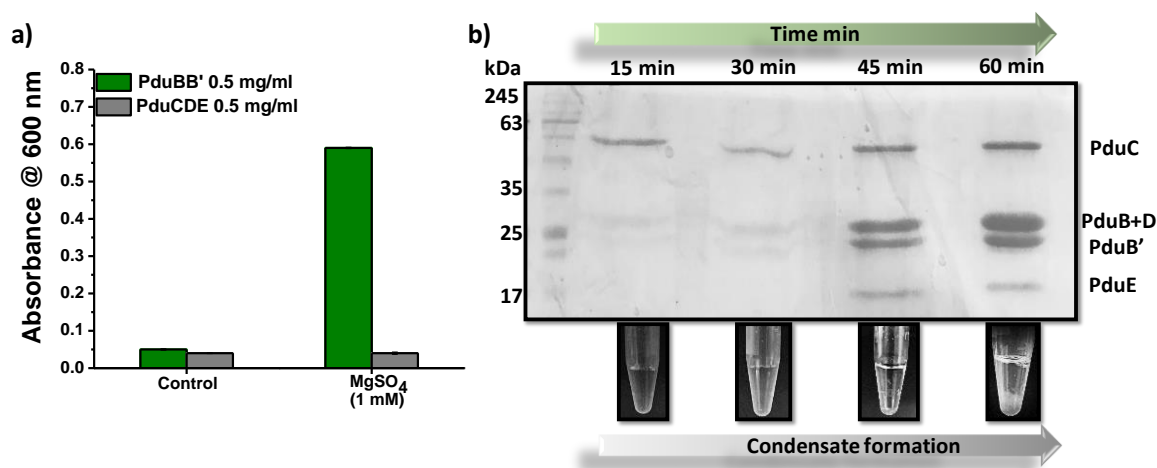
Figure 4.10: a) Turbidity assay of PduBB' (0.5 mg/ml) and PduCDE (0.5 mg/ml) in the absence and presence of 1 mM MgSO₄, b) Time dependant Co-condensation of PduBB' and PduCDE in solution.

Together, these results highlight probable role of Mg^{2+} ions in driving the shell-enzyme self-assembly and maintaining stability and optimum level of activity in purified PduMCP. Notably Mg^{2+} salt is an essential component of culture medium for producing PduMCP and its removal from the culture medium has detrimental effect on the formation of functional and intact PduMCPs in *Salmonella enterica* LT2 **(Fig. 4.11)**.

We induce PduMCP formation in *Salmonella enterica* LT2 strain in the absence and presence of MgSO₄. SDS PAGE of the cell lysates show that cells grown both in the presence and absence of MgSO₄ display the distinctive protein bands of PduMCP proteins **(Fig. 4.11a)**. This shows that *Salmonella enterica* LT2 produces the PduMCP component proteins whether MgSO₄ is present or not. The enzyme PduCDE is also generated in cells cultured both with and without MgSO₄. Diol dehydratase assay performed using crude cell lysate of *Salmonella enterica* LT2 further confirm that that PduCDE is generated in all cultures irrespective of the addition of Mg^{2+}**(Fig. 4.11b)**.

Notably, we detect higher concentration of propionaldehyde in lysate of *Salmonella enterica* LT2 grown in the absence of MgSO₄. This may be due to impaired PduMCP component protein self-assembly within cells in the absence of MgSO₄, resulting in a more accessible enzymatic core and higher propionaldehyde concentration[46, 55]. During our attempt to

purify PduMCP, the cells grown in the absence of $MgSO_4$ do not yield any detectable number of proteins as indicated by SDS-PAGE **(Fig. 4.11c)**, although the extract of the cells demonstrate the catalytic activity **(Fig. 4.11b)**. A possible explanation behind this observation could be the inability of the PduMCP proteins to assemble together in the absence of $MgSO_4$. As a result, the proteins would have failed to settle down during the centrifugation step of the purification protocol, thereby, lowering the yield of the PduMCPs. These experiments not only highlight the importance of Mg^{2+} salt in growth medium, but also hint towards the possible role of Mg^{2+} ions in production of functional and intact PduMCPs.
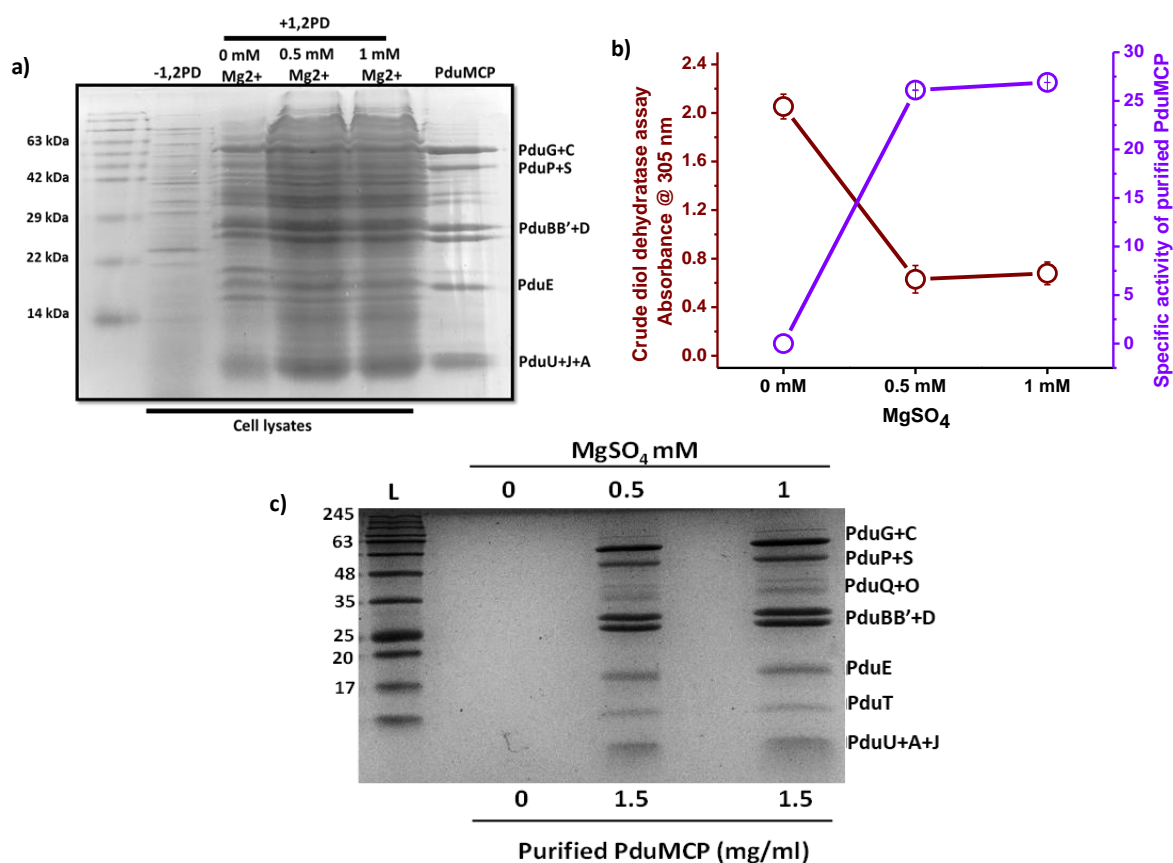
Figure 4.11: a) SDS PAGE of lysates of *Salmonella enterica* LT2 grown in the absence of 1,2-PD (lane 1), +1,2-PD and 0 mM $MgSO_4$ (lane 2), +1,2-PD and 0.5 mM $MgSO_4$ (lane 3), +1,2-PD and 1 mM $MgSO_4$ (lane 4); and purified PduMCP (lane 5), b) Crude diol dehydratase assay using cell lysate of *Salmonella enterica* LT2 grown in the absence and presence of 0.5 mM and 1 mM $MgSO_4$ (shown in maroon), Specific activity of purified PduMCP from *Salmonella enterica* LT2 grown in the absence and presence of 0.5 mM and 1 mM $MgSO_4$ (shown in purple), c) SDS PAGE of PduMCP purified from *Salmonella enterica LT2* grown in the absence and presence of 0.5 mM and 1 mM $MgSO_4$,

4.5 Conclusion

Understanding the self-assembly mechanism of MCPs remains a challenge because of the intrinsic small size of the bacterial cell and complex cellular organization. This calls for a minimalistic approach, where the factors controlling the self-assembly of the individual components are identified in isolation. The factors identified could be then explored and applied towards understanding the functioning of the entire microcompartments. In this regard, we attempt to probe the self-assembly behavior of a major shell protein (PduBB') and the signature enzyme (PduCDE) of PduMCP in solution phase. In the present work, we begin our study by looking at the effect of ionic strength of and macromolecular crowding of the surrounding environment on PduBB' assembly. We observe that the shell protein PduBB' separates as spherical droplets from the bulk solution displaying liquid like behavior. This is an interesting observation in the view of the recent reports where phase separation has been suggested to be the principle mechanism governing the self-assembly of different functional protein complexes at sub-cellular level, leading to protein aggregation as well as formation of membrane less organelles[27] [104] [105-107] [108, 109]. The concept of phase separation has been recently floated in the context of carboxysome and pyrenoid biogenesis[65, 66, 110]. Elegantly designed microscopic studies indicate the ability of carboxysome shell protein to undergo phase separation. For example, phase separation of, CsoS2, and its association with RuBisCo has been proposed to initiate α-carboxysome formation[66]. Similarly, phase separation of RuBisCo with shell protein CcmM drives the biogenesis of in case of ß-carboxysome[65]. These reports encourage us to explore the possible role of phase separation in the formation of several MCPs including PduMCP. Our findings show that PduBB' can experience phase separation under the right set of circumstances of molecule crowding and ionic strength. However, another major shell protein PduA fails to show phase separation behavior or recruit PduCDE enzyme under similar conditions **(Fig. 4.12)**. It is likely that PduBB', acts as a nucleating center during the early stage of PduMCP assembly. For this, PduBB' molecules need to come together in close vicinity and our study suggests that its phase transition governed by ionic strength, molecular crowding and strong shell-shell interaction might be the driving factor behind this phenomenon.

The combination of surface charge masking and kosmotropic effect[111] of metal ions appear to be the underlying cause of salt driven phase separation of shell protein PduBB'. In the absence of crowding agent, the shell protein PduBB' tends to self-associate in solution phase without undergoing any significant change in its secondary structure. The effect of salts on

the self-assembly of shell protein has been reported earlier, where the hexameric shell protein from *Haliungium Ocraceum* was found to self-associate with increase in ionic strength.[60]. The effect of divalent metal ions was found to be significantly higher than that of monovalent metal ions, an observation similar to the one made in the present study. We believe that ions present in the cellular environment may play crucial role in modulating the self-assembly of proteins of many different MCPs. It is therefore important that beside protein-protein interaction, the solution ionic strength is taken into consideration while discussing the self-assembly of various MCP proteins. Our *in vitro* data on phase separation of PduBB' points out at the significance of Mg^{2+} ions on shell protein assembly, and maintaining optimum stability of intact PduMCP. Further *in vivo* cellular studies would throw more light on this subject. However, this venture would require the use of high resolution imaging techniques that have recently gained attention for visualizing the phase separation of proteins in bacterial cells[112]. An alternate way would be to develop minimal compartments or cellular mimics and study the dynamics of the MCP proteins in a space and volume confined environment. Our previous work on the chaperone like behavior of PduBB' towards the signature enzyme[103], along with the present understanding about its phase separation behavior makes PduBB' a potential candidate for the engineering of synthetic nanoreactors.
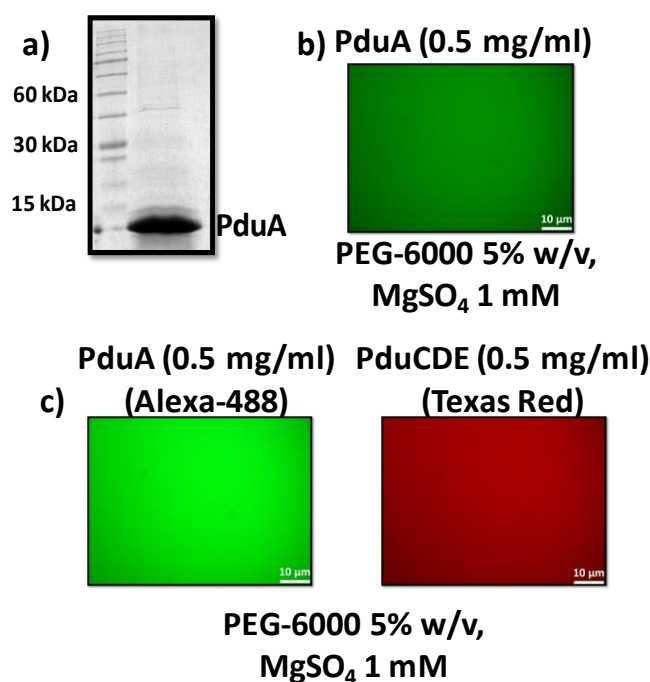
Figure 4.12: SDS PAGE of shell protein PduA (a), Microscopic visualization of (b) PduA (0.5 mg/ml) and (c) PduA: PduCDE mix (0.5 mg/ml each).

Additional note 4.1

Are metal ions an integral part of PduMCP?

*Given that metal ions, particularly Mg2+, have such a substantial impact on the stability and performance of PduMCP, it is unclear if these ions are actually present in PduMCP. In light of the findings that salts play a role in the self-assembly of PduMCP proteins and that their removal causes PduMCP destabilization, we investigated the elemental composition of purified PduMCPs from Salmonella enterica LT2. For this, we perform X-ray photoelectron spectroscopy (XPS) of purified PduMCPs after thorough washing of PduMCPs with de-ionized water (**see spectra in Box 4.2**). The survey scan of PduMCP shows peaks that correspond to different elements including as O, N, C, S, Na, Mg. Signals from elements such as O, N, C, S would arise as they are present in amino acids of proteins. Narrow scan between the range 1330 and 1300 eV confirms 1S peak signal for Mg (Fig. 5f). This data suggests the presence of ions such as Na^+, Mg^{2+} in PduMCP apart from O, N, C, and S. To eliminate the contribution of metal ions from the buffer, the PduMCPs are washed several times with de-ionized water and subjected to Inductively Coupled Plasma-Mass spectrometry (ICP-MS) analysis. ICP-MS analysis confirms the presence of Mg^{2+} ions in washed PduMCP sample (~13000 ppb). Mg^{2+} ions are not detected in the last supernatant wash sample indicating complete removal of free salts during the washing steps. Hence the Mg^{2+} detected are associated with PduMCPs. Further work would be needed to check if these ions become a part of PduMCPs during their biogenesis.*

XPS spectra showing 1S peak signal for Mg, survey scan (inset) showing signal for elemental composition (S, C, N, O, Na, Mg) in PduMCP.
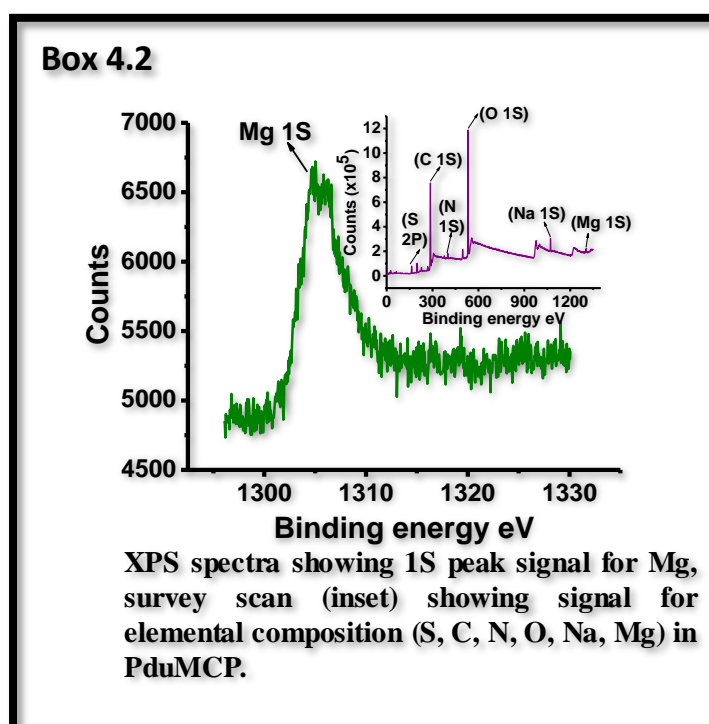
Chapter 5

Distribution and Functions of Intrinsically Disordered Regions in 1,2-Propanediol Utilization Microcompartment

5.1 Introduction

Intrinsically disordered regions or IDRs defy Anfinsen's sequence-structure-function postulate and can take on a variety of conformations depending on cellular requirements. IDRs do not undergo co-operative folding and are unique in their ability to adapt for various biological functions by binding with multiple cellular partners due to their high degree of flexibility [113] [114]. Beside participating in interactions, IDRs have also been shown in recent years to facilitate the formation of membrane-less organelles such as nucleolus [115], RNP granules [116], P-bodies [117], and so on. Weak multivalent interactions between IDRs have been proposed to be critical for phase separation and the formation of membrane-less organelles [118]. Identification of the IDRs in complex multi-protein compartments is therefore essential to enhance our understanding about how IDRs function in subcellular compartmentalization.

Our current understanding of MCP biogenesis suggests that thousands of protein subunits self-assemble via shell protein-enzyme interaction [47]. The role of disordered regions in the self-assembly and formation of carboxysomes have recently been reported in the literature [65, 66, 119]. However, the distribution and potential roles of the disordered regions in metabolosomes are not well-understood. Due to the complexity and multicomponent makeup, PduMCPs are an interesting paradigm to study and examine for the distribution of IDRs among their shell proteins and enzymes. In previous chapters, we learnt that the outer shell of PduMCP protects the luminal enzymes from thermal denaturation. The major shell protein PduBB' interacts with the signature enzyme diol dehydratase PduCDE, preserving its native conformation under thermal stress [67]. The self-assembly of PduBB' can be fine-tuned by optimizing ionic strength and molecular crowding of the surrounding environment. PduBB' displays phase separation behaviour and has the ability to co-phase separate PduCDE, improving its activity.

In the present chapter we concede the fact that encapsulation of thousands of copies of enzyme molecules, with a high likelihood of shell-enzyme interactions, would necessitate a considerable deal of structural flexibility. Thus it's likely that, in addition to structured domains, disordered regions in shell proteins and enzymes aid in outfitting the structure of PduMCPs. In the present work, we use bioinformatics analysis to identify the probable disordered regions in the shell proteins and enzymes of PduMCP. Using a combination of computational and biochemical technique, we assess the potential roles of the disordered

regions in the structural flexibility and solvent accessibility of the major shell protein PduBB'. We further use computational simulation to predict the nature and strength of association between shell protein PduBB' and enzyme PduCDE. Our results indicate the presence of IDRs in both shell proteins and enzymes and may be responsible for dynamic protein-protein interactions.

5.2 Prevalence of Intrinsically disordered regions in Pdu proteins

The distribution of disordered regions among the shell proteins and enzymes of PduMCP is determined using PONDR server which gives percentage of residues that fall in the disordered regions **(Table 5.1)**. Among the shell proteins, PduK and PduN are predicted to have the highest composition of disordered regions with ~46% of their residues falling in the regions that do not fold into a secondary structure **(Fig. 5.1a)**. They are followed by PduB (25.27%), PduJ (22.96%) and PduA (17.02%). The disordered regions are mainly localized in the terminal regions (N and C) of these shell proteins **(Fig. 5.1b-h and Fig. 5.2)**. For instance, the C-terminal extension of PduK is predicted to be intrinsically disordered from A94-A160 residues. Notably, this region includes a [Fe-S] center (CNLCLDPKCPRQKGEPRTLC) **(Fig. 5.1c)**, that may be involved in electron transfer reactions [120]. The significance of the C-terminal region of PduK in the context of protein-protein interaction is however, not well understood. We identify molecular recognition features in the region Val130-Leu136 as predicted by MoRFPred server [76]. MoRFPred helps in identification of residues within the disordered regions that may undergo disorder to order transition during interactions. The region Val130-Leu136 partly overlaps with the [Fe-S] centre and includes two cysteine residues, indicating that the disordered C-terminal region of PduK may participate in iron binding. Both PduA and PduJ have short disordered segments in their C-terminals. The segments L88-Q94 and I76-A91 in PduA and PduJ respectively have high disorder propensity. This is more apparent from the crystal structures of PduA (PDB id: 3NGK) and edge mutant PduJ (K26A, PDB id: 5D6V). that shows coil-helix-coil endings I76-Q94 and I76-A91 in PduA and PduJ (K26A) respectively **(Fig. 5.3a, i-ii)**. In case of PduB the 37 amino acids N-terminal extension (M1-A37) has high disorder propensity **(Fig. 5.1f)** with ~50% of its residues exhibiting high disorder propensity **(Fig. 5.1a and Fig. 5.1g)**.

Table 5.1: Percentage of disordered regions in shell protein and enzymes of PduMCP			
Pdu-MCP Proteins	Function	No. of Residues	% disorder PONDR-VSL2
PduA	Shell protein	94	17.02
PduB	Shell protein	270	22.96
PduB'	Shell protein	233	12.02
PduC	Diol dehydratase subunit	554	20.94
PduD	Diol dehydratase subunit	224	37.95
PduE	Diol dehydratase subunit	173	32.95
PduG	Reactivase	610	5.25
PduH	Reactivase	116	19.83
PduJ	Shell protein	91	25.27
PduK	Shell protein	160	46.88
PduL	Phosphotransacylase	210	15.71
PduM	Shell protein	163	15.34
PduN	Shell protein	91	46.15
PduO	Adenoacyltransferase	336	22.92
PduP	Propionaldehyde dehydrogenase	464	27.80
PduQ	Propanol dehydrogenase	370	8.92
PduS	B12 reductase	451	13.53
PduT	Shell protein	184	12.50
PduU	Shell protein	116	11.21
PduV	Unknown function	150	12.67
PduW	Propionate kinase	404	21.53
PduX	L-threonine kinase	300	13.00
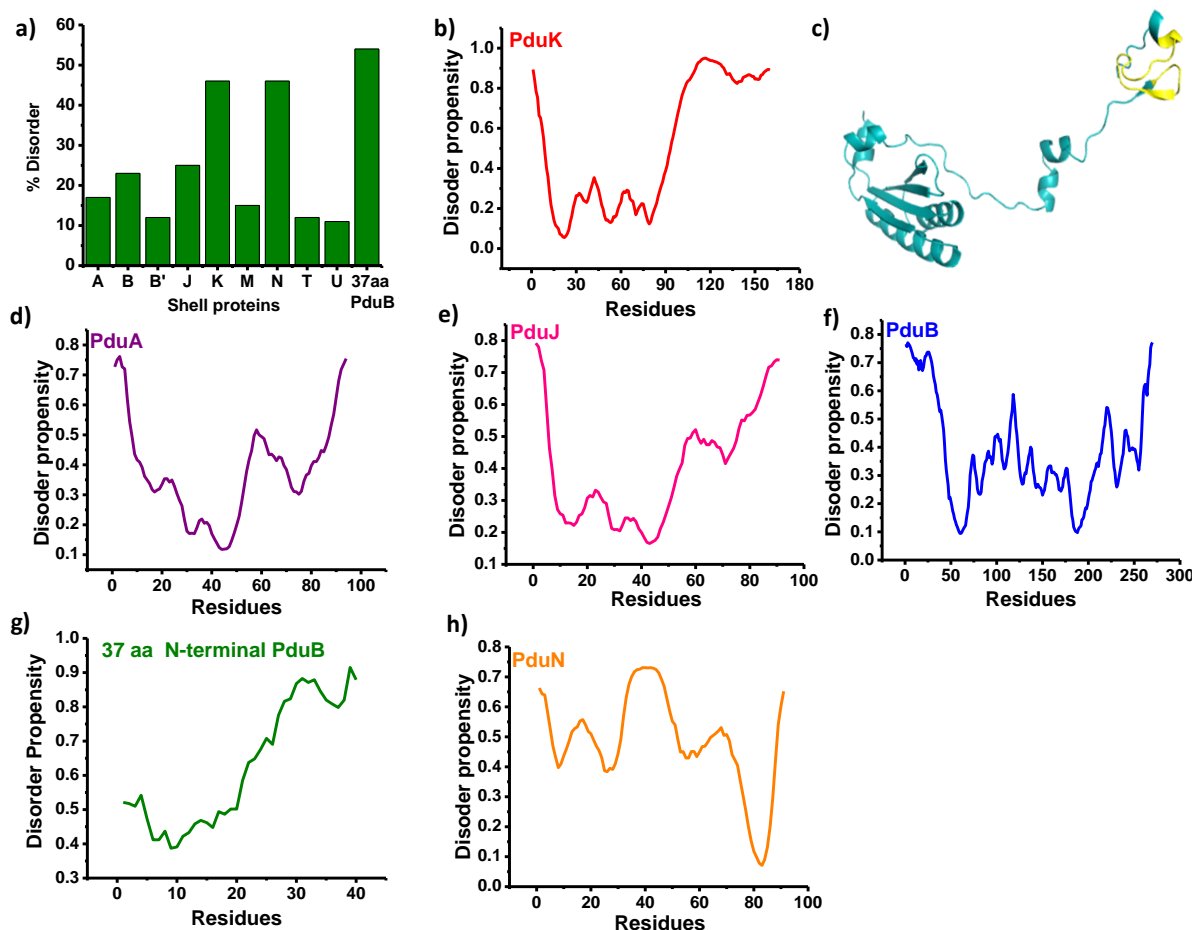
Fig 5.1: a) Disorder percentage among the shell proteins of PduMCP and 37 amino acids N-terminal region of PduB; b) disorder propensity of PduK determined using PONDR web server, c) Stucture of PduK monomer modeled using Galaxy TBM server. Cysteine rich region (NLCLDPKCPRQKGEPRTLC) has bee marked in yellow, disorder propensity of d) PduA, e) PduJ, f) PduB, g) 37 amino acids N-terminal region of PduB and h) PduN determined using PONDR web server. A propensity score more than 0.5 indicates high probability of disorder.

PduK

```
1      MANKEHRVKQ SLGLLEVCGL ALAISCADIM AKSASITLLA LEKTNGSGWM
VSL2   DDDDDDDD
```

```
51     VIKITGDVAS VQAAITTGAH FAEQWNGLVA HKVIARPGEG ILLAETPSPS
VSL2                                                  DDDDDDD
```

```
101    VIEPEPEASE IADVVSEAPA EEAPQESELV SCNLCLDPKC PRQKGEPRTL
VSL2   DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD
```

```
151    CIHSGKRGEA
VSL2   DDDDDDDDDD
```

PduN

```
1      MHLARVTGAV VSTQKSPSLI GKKLLLVRRV SADGELPASP TSGDEVAVDS
VSL2   DDDDD      DDDDDDDD             DDDDDDDD  DDDDDDDDDD
```

```
51     VGAGVGELVL LSGGSSARHV FSGPNEAIDL AVVGIVDTLS C
VSL2   D                DDDDDD                  DD D
```

PduB

```
1      MSSNELVEQI MAQVIARVAT PEQQAIPGQP QPIRETAMAE KSCSLTEFVG
VSL2   DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD
```

```
51     TAIGDTLGLV IANVDTALLD AMKLEKRYRS IGILGARTGA GPHIMAADEA
VSL2
```

```
101    VKATNTEVVS IELPRDTKGG AGHGSLIILG GNDVSDVKRG IEVALKELDR
VSL2                  DDDDD
```

```
151    TFGDVYGNEA GHIELQYTAR ASYALEKAFG APIGRACGII VGAPASVGVL
VSL2
```

```
201    MADTALKSAN VEVVAYSSPA HGTSFSNEAI LVISGDSGAV RQAVTSAREI
VSL2                  DD DDD
```

```
251    GKTVLATLGS EPKNDRPSYI
VSL2            DD DDDDDDDDDD
```

PduJ

```
1      MNNALGLVET KGLVGAIEAA DAMVKSANVQ LVGYEKIGSG LVTVMVRGDV
VSL2   DDDDD
```

```
51     GAVKAAVDAG SAAASVVGEV KSCHVIPRPH SDVEAILPKS A
VSL2           DD                 DDDD DDDDDDDDDD D
```

PduA

```
1      MQQEALGMVE TKGLTAAIEA ADAMVKSANV MLVGYEKIGS GLVTVIVRGD
VSL2   DDDDDDD
```

```
51     VGAVKAATDA GAAAARNVGE VKAVHVIPRP HTDVEKILPK GISQ
VSL2           DD                      DDD DDDD
```

Fig. 5.2: Disorder propensity of shell proteins PduK, PduB, PduA, PduJ and PduN determined using PONDR web server.

5.3 Vertex shell protein of PduMCP has high backbone flexibility

PduN, which forms pentameric assemblies have disordered segments distributed all throughout the protein **(Fig. 5.1h)**. Homology modelling of PduN also confirms this prediction and shows that nearly 50% of its amino acids (segments A9-L24, R29-G43, V48-G56 and G63-L80) fall in the disordered regions **(Fig. 5.3a, iii)**. Disorder in a significant portion of shell protein PduN render the shell protein PduN highly flexible. Using CABS Flex simulation of modelled PduN monomer (equivalent to 10 ns MD simulation) we evaluate the flexibility of PduN. The simulation result shows that PduN has four regions with high mean residual fluctuation or rmsf values **(Fig. 5.3b, i-ii)**. These regions with high rmsf values overlap with the disordered regions in the protein structure **(Fig. 5.3a, iii)**. High residual fluctuations confirms that the disordered regions in PduN are highly flexible. Higher backbone flexibility in PduN may be the reason why purified shell protein PduN display random coil property as measured by CD spectroscopy in our another study [61]. Increased backbone flexibility may also help PduN shell protein to occupy the vertices of the PduMCPs and increase its potency to interact with multiple shell protein partners.
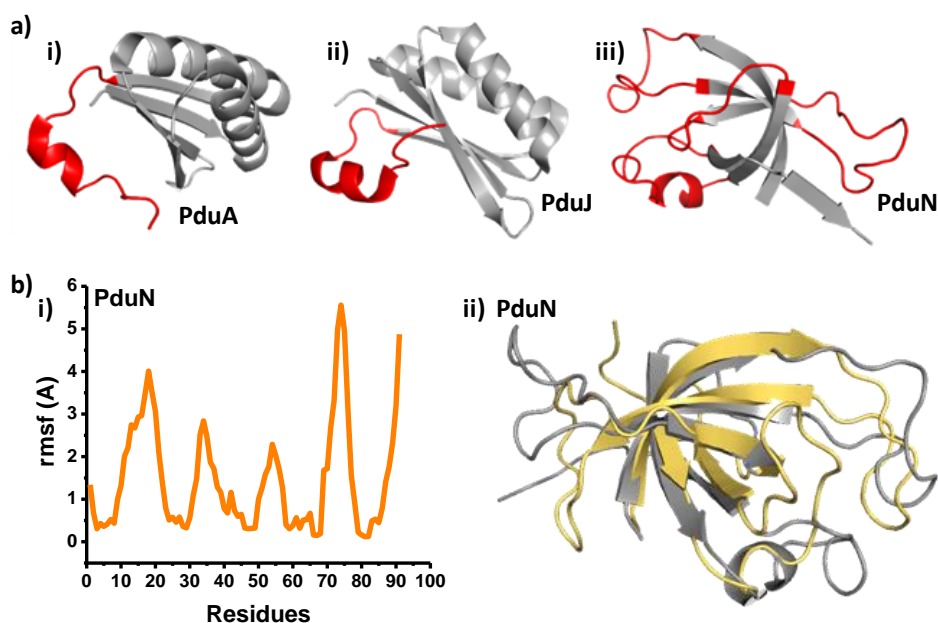
Fig. 5.3: a) Structures of shell proteins (i) PduA, (ii) PduJK26A and (iii) PduN with identified disordered regions highlighted in red.   a) Rresidual flexibility plot of PduN determined using CABS Flex web server, b) PDB structures of PduN before (in grey) and after (in yellow) simulation run.

5.4 The signature enzyme of PduMCP has disordered N-terminal extensions

Among the enzymes, middle and the smaller subunits of diol dehydratase PduCDE have the highest disorder percentage. Subunit PduD has more than 35% of its residue in the disordered region. PduD is followed by PduE with ~32% of its residues in the disordered region. Like the shell proteins, the N-terminal regions of both PduD and PduE have high disorder propensity **(Fig. 5.4, a-b)**. Since the crystal structure of PduCDE is not available in the literature, we perform homology modelling of the all the three subunits of the enzyme. While the catalytic subunit PduC is completely modelled, the N-terminal segments of PduD (M1-G46) and PduE (M1-R36) are eliminated in the model **(Fig. 5.4c)**. In the crystal structure of a PduCDE homologue (diol dehydratase, PDB id: 1DIO) also the N-terminal regions of subunit D and subunit E could not be crystallized. This could be because of the disorder in the N-terminal regions of this protein that fail to get crystallized due to lower electron density. In the following section, we discuss the probable roles of these disordered regions in Pdu proteins.
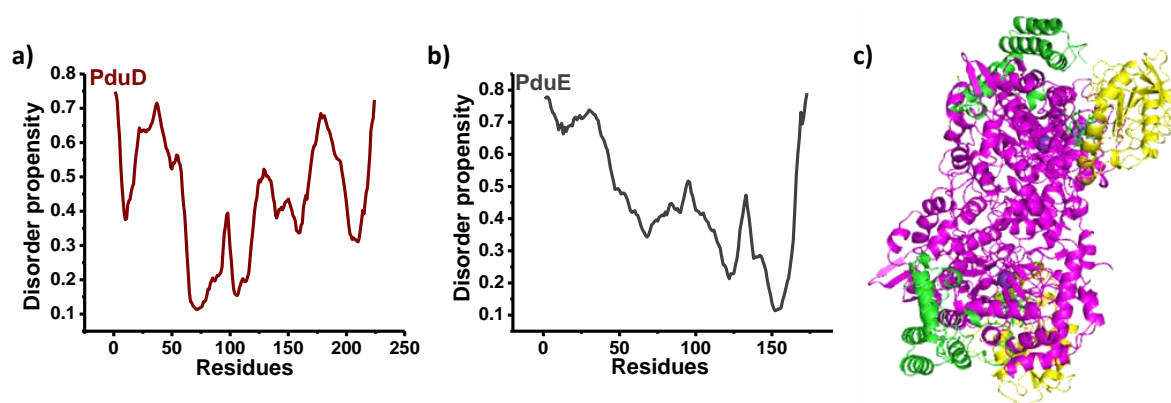
Fig. 5.4: a) Residual disorder propensity of PduD, b) Residual disorder propensity of PduE, c) Homology model of PduCDE using crystal structure of diol dehydratase from *Klebsiella oxytoca* (PDB ID: 1 DIO), as template

5.5 Flexibility in the N-terminal region of a major shell protein PduBB'

PduBB' is a combination of two proteins, PduB and PduB'. A single gene with two start codons is responsible for encoding both proteins. PduB, which has additional 37 amino acids N-terminal extension is longer than PduB' [67]. When monomeric subunit of PduB is subjected to simulation for 10 ns using CABS flex simulation server, the terminal region of the protein exhibits ~10 fold higher rmsf value than its structured domain **(Fig. 5.5)**. This

indicates that the N-terminal region of PduB has a flexibility, which may be attributed to the disorder in this region of the protein.
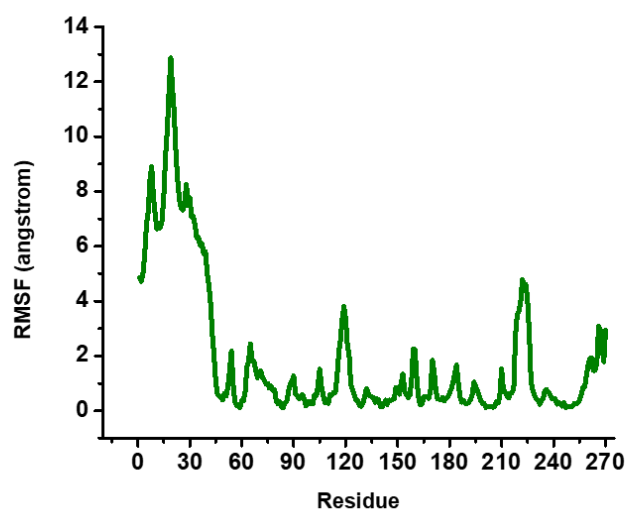
Fig. 5.5: Residual flexibility plot of PduB determined using CABS Flex web server

Higher flexibility regions in proteins have high solvent accessibility and more prone to proteolytic cleavage. To check if the N-terminal region of PduBB' is more susceptible to proteolytic digestion, we perform limited proteolysis of the shell protein using trypsin followed by visualization of the protein sample on SDS PAGE gel. In the absence of any proteolytic enzyme, PduBB' gives two distinct bands corresponding to its constituent proteins (PduB, 270 residues and PduB', 233 residues). Treating with trypsin for a short duration of 0.5 min, we observe decrease in the band intensity corresponding to PduB (270 residues) **(Fig. 5.6a)**. Prolonging the trypsin digestion time results in further decrease in the band intensity for PduB and we do not see any PduB band after 5 min of tryptic digestion. The decrease in the band intensity for PduB may be attributed to the cleavage of the N-terminal region of the shell protein. This would result in the increase in the band intensity for PduB' at 0.5 min sample. However, further decrease in the PduB band intensity does not bring about simultaneous increase in the PduB' band intensity. To understand the underlying reason behind this observation, we perform limited proteolysis of the component shell proteins of PduBB' (i.e PduB and PduB'), which has been described in the following section.

5.6 Disordered N-terminal region of PduBB' provides solubility to the shell protein

To get PduB, we generated PduBM38L mutant that gives a single protein band on SDS PAGE. Treating PduBM38L with trypsin for a short duration of 0.5 min gives two protein

bands corresponding to PduB and PduB' (**Fig. 5.6b**). With the increase in the proteolysis time, the band intensities for PduB and PduB' decreases and increases respectively. This confirms that the extended N-terminal region of PduB is susceptible to proteolytic cleavage which is cleaved off during the tryptic digestion. Here it is important to mention that band intensity of PduB' never approaches the initial band intensity of PduB. One may anticipate that over time, the intensity of PduB' will equal the initial band intensity of PduBM38L if just the N-terminal portion of PduB is being cleaved by trypsin. There are two explanations that could explain this finding. First, other regions in the structured domains of both PduB and PduB' experience proteolytic cleavage in addition to the extended N-terminal. Second, only the longer form of the shell protein (PduB) undergoes proteolysis (both in the N-terminal and structured domain) and the truncated version of the protein (PduB') formed has lower susceptibility towards proteolysis. In order to deduce the correct explanation, we subject the truncated version of PduBB', (PduB') to limited proteolysis using trypsin. As shown in **Fig 5.6c**, there is no significant change in the band intensity for PduB' in the absence and presence of trypsin. This observation suggests that shell protein PduB does not undergo proteolytic digestion within a time period of 60 min. This may explain why we do not observe any significant change in PduB band intensity during proteolysis of PduBB'. This also suggests that it is the longer form of the shell protein that undergoes proteolytic digestion upon increasing the time of trypsin treatment. Difference in the susceptibility of the two forms of the shell protein (PduB and PduB') may be attributed to the difference in their solubility. Upon purification, both PduBB' and PduBM38L remains soluble. The truncated protein PduB' tend to form associates giving a turbid solution as reported in our earlier work [67]. The association of PduB proteins in solution would reduce the accessibility of its residues to the catalytic site of the proteolytic enzyme. Together, this study reveals that the N-terminal region of PduBB' renders solubility to the shell protein [67].
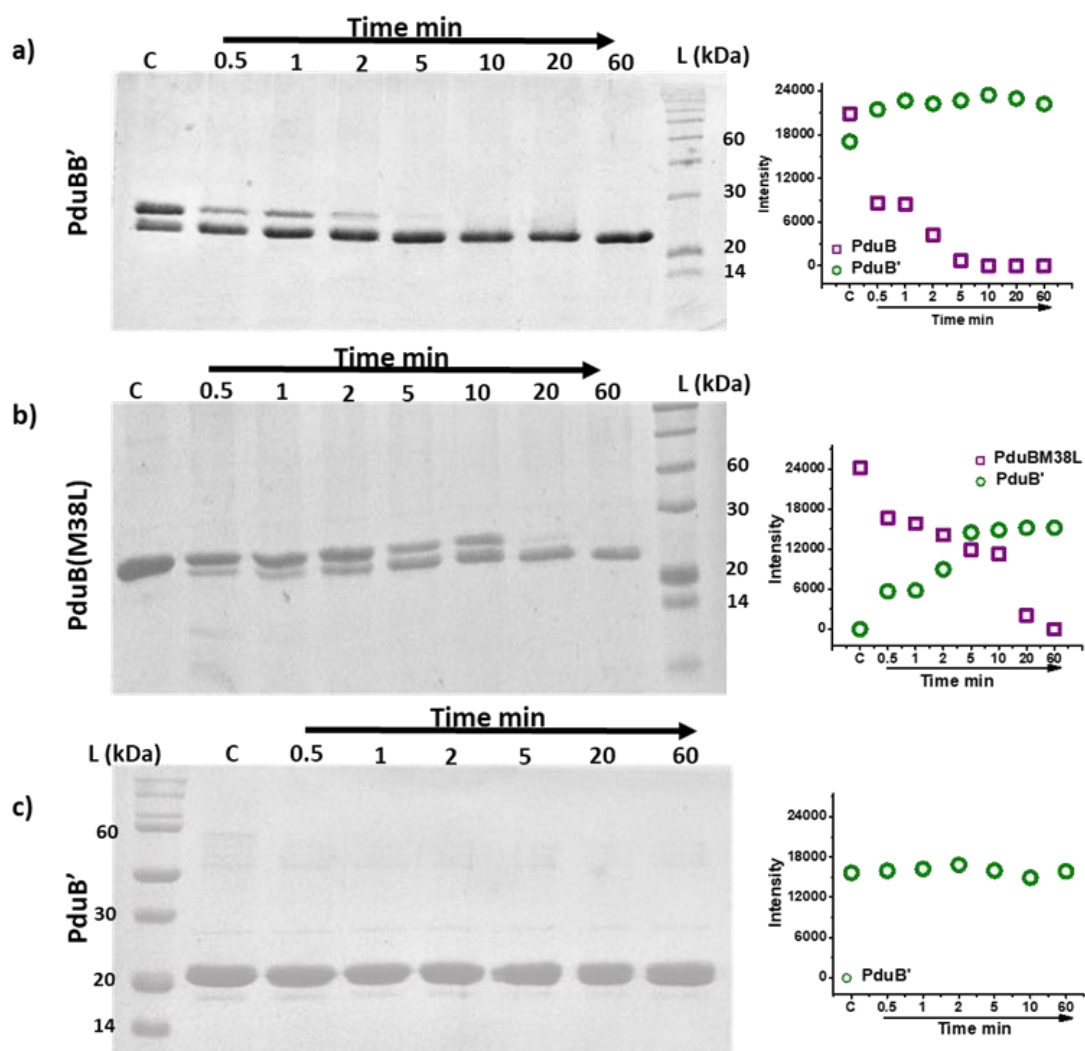
Fig 5.6: a) Limited proteolysis of shell proteins a) PduBB', b) PduBM38L), c) PduB'. The panel on the right shows change in protein band intensities with time of trypsinization.

5.7 Disordered regions and shell-enzyme interaction

The terminal regions of the shell proteins have been suggested to participate in binding enzymes and facilitating their encapsulation within the MCP [47, 79, 121]. For instance, the N-terminal of PduB has been shown to link the enzyme core of the PduMCP to its outer shell [47]. Earlier, the N-terminal region of PduD was shown to target PduCDE within PduMCP [79]. Our earlier report has shown that PduBB' has micromolar affinity towards PduCDE and conserves its catalytic property under heat stress. It is very likely that the major shell protein PduBB' helps the enzyme PduCDE to get encapsulated within PduMCP. To understand how disordered terminal regions may contribute to PduB-PduD interaction, we use IUPred server

and its ANCHOR algorithm to identify residues that are part of the disordered as well as interacting regions in PduB and PduD **(Fig. 5.7)**. In case of PduB, the first 16 residues of the N-terminal are involved in binding and this binding region is followed by a proline-glutamine rich disordered region (R17-A37) **(Fig. 5.7a, i-ii)**. Modelling of the 37 amino acids of PduB using Galaxy server [69], also shows a helical region followed by a disordered region **(Fig. 5.7a, iii)**. Similar observation is made in case of PduD, where first 20 residues (M1-K20) in the N-terminal is identified as a binding region. This is followed by an intrinsically disordered region from G21-G44 **(Fig. 5.7b, i-ii)**. Modelling of the first 44 amino acids of PduD using Galaxy server shows a helical region followed by a disordered region **(Fig. 5.7b, iii)**. The results of molecular recognition features (MoRFs) analysis also demonstrate the presence of molecular recognition features in the predicted helical region of N-terminal segments of PduB (Q9 IMAQVI A16) and PduD (R9 QIIED V15). The predicted MoRF of PduB includes I10, that has earlier been shown to be important for the encapsulation enzymes within PduMCP [47]. In the likely event that the terminal regions of shell protein and enzyme interact, the disordered segment in the N-terminal region of these proteins would offer the required conformational flexibility and make it easier for the preceding helical region to bind with the protein partners.

a)

Shell protein PduB

Protein binding   **Disordered**

i) MSSNELVEQI MAQVIARVAT PEQQAIPGQP QPIRETAMAE KSCSLTEFVG TAIGDTLGLV
IANVDTALLD AMKLEKRYRS IGILGARTGA GPHIMAADEA VKATNTEVVS IELPRDTKGG
AGHGSLIILG GNDVSDVKRG IEVALKELDR TFGDVYGNEA GHIELQYTAR ASYALEKAFG
APIGRACGII VGAPASVGVL MADTALKSAN VEVVAYSSPA HGTSFSNEAI LVISGDSGAV
RQAVTSAREI GKTVLATLGS EPKNDRPSYI

ii)

iii)

b)

Subunit PduD

Protein binding   **Disordered**

i) MEINEKLLRQ IIEDVLRDMK GSDKPVSFNA PAASTAPQTA APAGDGFLTE VGEARQGTQQ
DEVIIAVGPA FGLAQTVNIV GLPHKSILRE VIAGIEEEGI KARVIRCFKS SDVAFVAVEG
NRLSGSGISI GIQSKGTTVI HQQGLPPLSN LELFPQAPLL TLETYRQIGK NAARYAKRES
PQPVPTLNDQ MARPKYQAKS AILHIKETKY VVTGKNPQEL RVAL

ii)

iii)

Fig 5.7 a) Amino acid sequence of PduB (i), Disorder propensity (in pink) and binding region (in blue) identified using IUPred and ANCHOR webserver (ii), Model of PduB N-terminal showing Disordered (in pink) and binding region (in blue) identified using IUPred and ANCHOR webserver (iii), b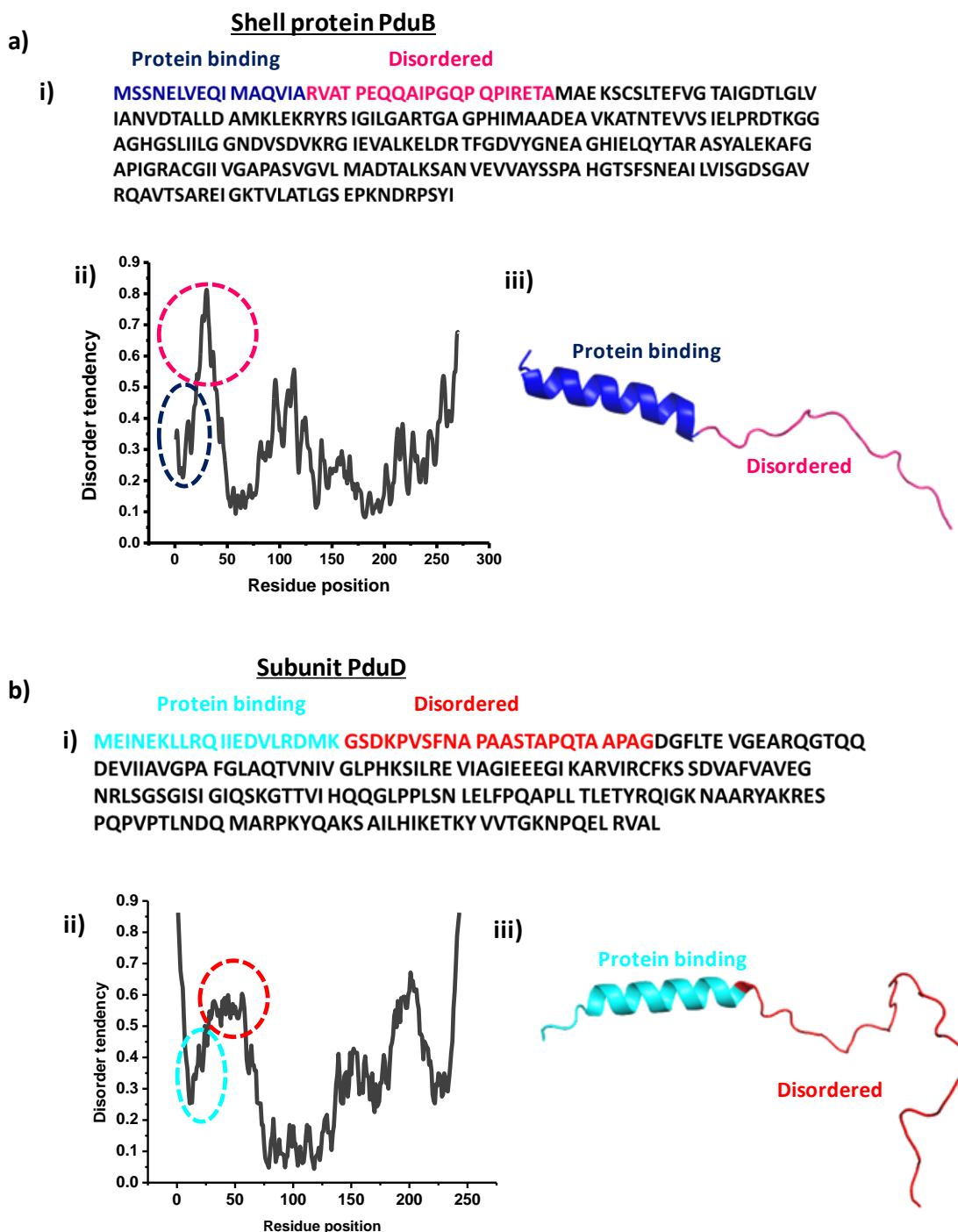) Amino acid sequence of PduD (i), Disorder propensity (in pink) and binding region (in blue) identified using IUPred and ANCHOR webserver (ii), Model of PduD N-terminal showing Disordered (in pink) and binding region (in blue) identified using IUPred and ANCHOR webserver (iii). Modelling of N-terminal regions of PduB and PduB was performed using Galaxy TBM server

Flexible docking between the N-terminal segments of PduB and PduD using CABS dock server 8 possible modes of binding between two terminal regions with a mean binding free energy (ΔG) of -9.2 kcal/mol **(Fig. 5.8)**. The presence of MoRFs within a terminal disordered region may not be limited to MCP proteins but may have role in other bacterial proteins. For example, polar organizing protein (PopZ) in *Caulobacter crescentus*, has a disordered N-terminal region that binds to its partner proteins. The interaction is mediated by a short helix of ~13 amino acids (MoRFs) within the disordered region of the protein [122].
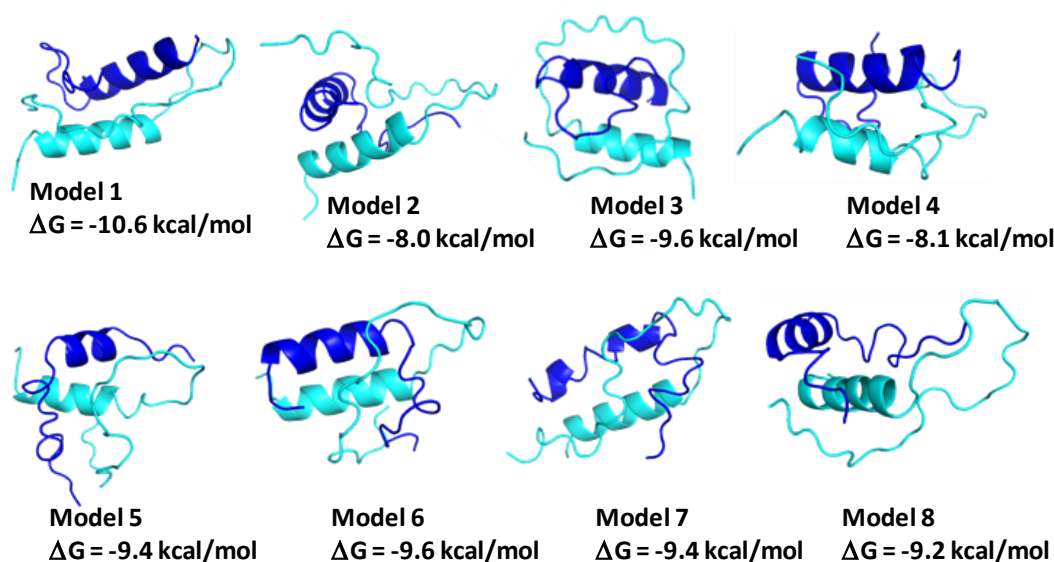
Fig. 5.8: Docked models showing interaction between N-terminal regions of PduB and PduD.

5.8 N-terminal regions of PduB and PduD mediate flexible shell-enzyme association

The importance of the terminal regions of shell protein and enzyme in MCP formation has been well explored in literature. In the context of PduMCP, the N-terminal regions of middle subunit of PduCDE (PduD) and PduBB' have shown to be essential for the formation of a fully intact PduMCP.[47, 79] Our findings suggest the role of 37 amino acid N-terminal region of PduBB' in providing solubility to the shell protein and regulating its self-assembly behavior in solution phase. In order to understand the significance of N-terminal region of PduD subunit of PduCDE, we delete R2-K20 residues of the N-terminal region of PduD in PduCDE that has been proposed to be involved in mediating enzyme encapsulation **Fig. 5.9a**. Using time resolved fluorescence anisotropy we determine the rotational correlation time of Acrylodan labeled PduBB' in the absence and presence of PduC(Δ2-20D)E. The rotational

co- relational time of PduBB' does not change significantly in the presence of PduC(Δ2-20D)E, indicating weaker interaction between the two proteins **(Fig. 5.9b)**. This result is consistent with a previous report that emphasized the importance of the N-terminus of PduD in the packaging enzyme within PduMCP.[79]

To gain an insight into our experimental data, we perform flexible docking of N-terminal sequence of PduB with full length PduD and PduD (Δ1-20). For flexible docking, CABS Dock server[71] is used and *de novo* modeled PduD along with N-terminal sequence of PduB are uploaded as input files. During the simulation run, the N-terminal sequence of PduB takes the helical conformation and interacts with PduD (ΔG = -15.3 kcal/mol) **(Fig. 5.9c and Table 5.2)**. Beside 3 hydrogen bonds and 2 salt bridges, hydrophobic contacts are shown to link the terminal region of PduB with PduD. The residue Ile10 of PduB N-terminal makes hydrophobic contacts with the hydrophobic residues in the N-terminal region of PduD. Notably, Ile10 has been shown earlier to play crucial role in enzyme encapsulation within PduMCP.[47] In case of PduD (Δ1-20), the N-terminal region of PduB interacts with the protein with a significantly lower affinity (ΔG = -4.1 kcal/mol), mediated by hydrophobic interactions **(Fig. 5.9d and Table 5.3)**. This suggests that the absence of the N-terminal region of PduD may weaken the association of the enzyme with the shell protein.

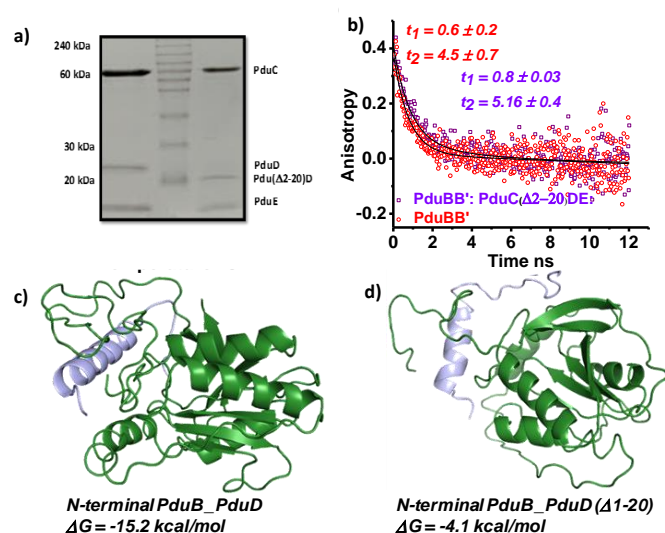**Fig. 5.9: a) SDS PAGE of full length PduCDE and N-terminal truncated PduCDE [PduC(Δ2-20D)E]b) Fluorescence anisotropy of PduBB' in the absence and presence of PduC(Δ2-20D)E, c) Flexible docking of N-terminal sequence of PduB with *de novo* modeled PduD using CABS dock server, d) Flexible docking of N-terminal sequence of PduB with *de novo* modeled PduD (Δ1-20) using CABS dock server.**

In order to understand the nature of interaction between the N-terminal regions of PduB and PduD, we use a docked model of PduB and PduD and subject it to MD simulation for 20 ns (**Fig. 5.10, a-d**). The extended N-terminal regions of PduB and PduD provide a very flexible association of the two proteins (**Fig. 5.10a, i**). This is evident from the RMSF plots of the two proteins that show a very high flexibility in the N-terminal region of PduB and PduD as compared to their structured domains (**Fig. 5.10a, ii and iii**). During the simulation run, 8 possible salts bridge interactions participate in holding the PduB-PduD complex. Due to a high flexibility in the terminal regions of the two proteins, the salt bridges form and break at different time points, as shown in the residual distance v/s time plot (**Fig. 5.10c**). In case of N-terminal truncated PduD (Δ N-terminal PduD), the N-terminal of PduB binds with the structured domain of PduD (**Fig. 5.10b, i**). During the simulation run, the interaction between the PduB and Δ N-terminal PduD remains comparatively more rigid with low RMSF values (**Fig. 5.10b, ii and iii**).

Table 5.2: Residual contacts between PduD and N-terminal region of PduB			
PduD	PduB	Type of bond	Distance Å
Ala 32	Gln 9	H-bond	3.1
Asp 45	Gln 23	H-bond	2.8
Ser 2	Glu 2	H-bond	2.8
Lys 24	Glu 5	Salt bridge	2.8
Asp 14	Arg 17	Salt bridge	2.6
Val 15	Ile 10	Hydrophobic	2.9
Leu 16	Ile 10	Hydrophobic	3.5
Phe 28	Leu 6	Hydrophobic	3.2
Table 5.3: Residual contacts between PduD (Δ1-20) and N-terminal region of PduB			
PduD	PduB	Type of bond	Distance Å
Leu 73	Ala 25	Hydrophobic	2.7
Ile 79	Val 7	Hydrophobic	2.9
Val 80	Val 7	Hydrophobic	3.5

Due to reduced flexibility, only two possible salts bridges mediate the interaction between the two proteins (**Fig. 5.10d**). This is indeed an important observation, as it highlights the significance of the terminal regions in providing a flexible interaction between the shell protein and enzyme. It is very likely that a flexible association may increase the chance of residues in the terminal regions of the shell protein and enzyme to interact with one another. This could be the reason behind the involvement of 8 salt bridges in PduB-PduD interaction as compared to 2 salt bridges in PduB-(Δ N-terminal PduD), over a period of time. Participation of multiple residues may contribute towards stronger binding between the shell protein and enzyme when the terminal regions are present. This is evident from the binding

free energies of the two complexes determined using Hawkdock server [123]. In case of the PduB-PduD complex the binding, free energy is -94.8 kcal/mol, compared to -58.1 kcal/mol in case of PduB-Δ N-terminal PduD complex, at the end of simulation run. This observation is in line with our previous report where we showed that deletion of N-terminal region PduD weakens the binding of PduCDE with the shell protein PduBB' [67]. The N-terminal modeling, MoRF prediction and MD simulation taken together suggest that the shell protein and enzyme would initially bind through MoRFs in their N-terminal regions. Over time, however, the link between the two proteins would be strengthened by numerous salt bridges. But because the N-terminal regions of shell proteins and enzymes are flexible, the shell protein-enzyme association would be exceedingly dynamic.
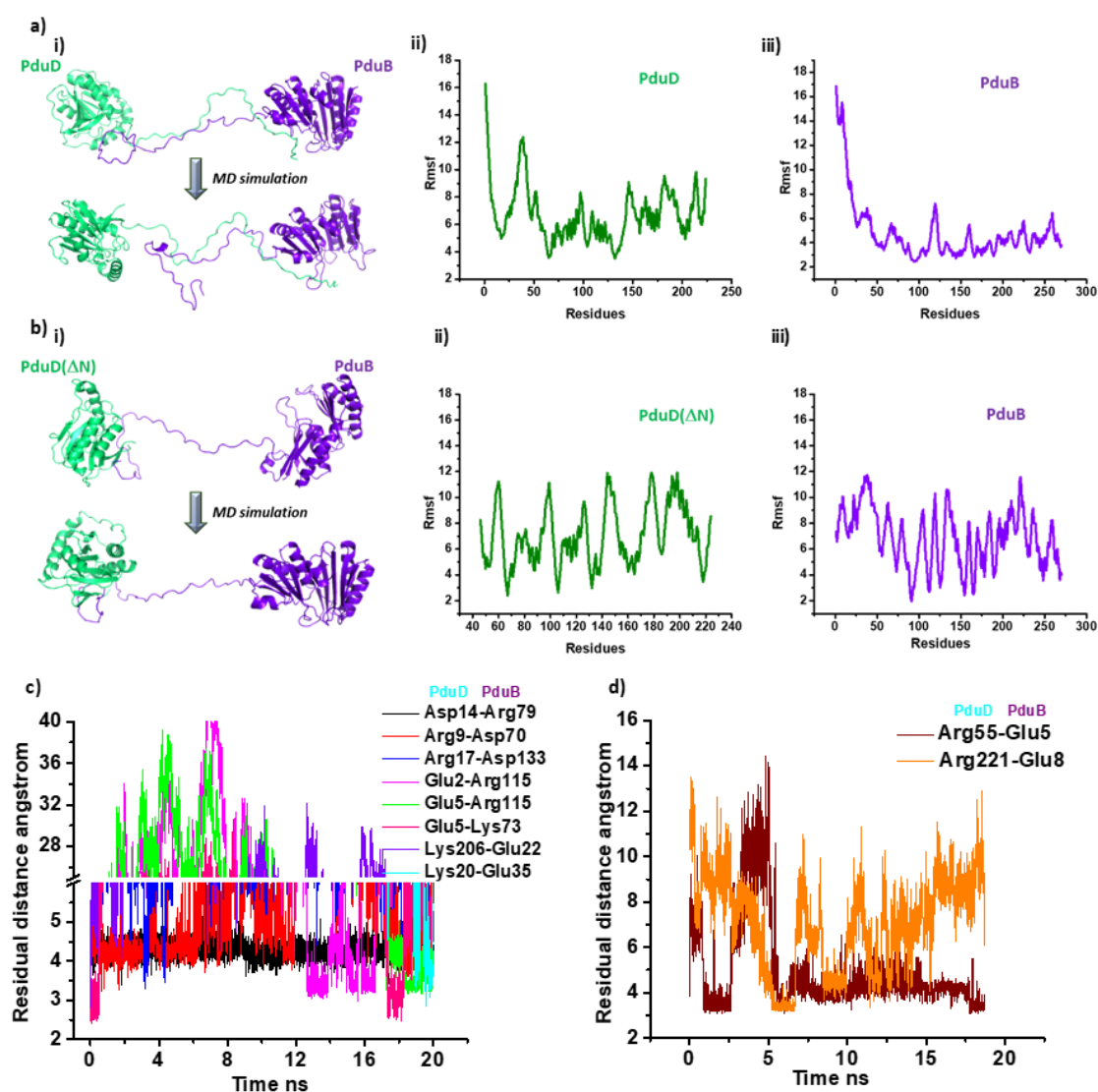

Fig 5.10: a) ClusPro docked model of PduB and PduD pre and post MD simulation (i), RMSF plots showing residual flexibility of PduD (ii) and PduB (iii) during the simulation; b) Docked model of PduB and Δ N-terminal PduD pre and post MD

simulation (i), RMSF plots showing residual flexibility of Δ N-terminal PduD (ii) and PduB (iii) during the simulation; c) Plot of residual distance v/s time showing change in the distance between the charged amino acids of PduB and PduD during the simulation run; d) Plot of residual distance v/s time showing change in the distance between the charged amino acids of PduB and Δ N-terminal PduD during the simulation run.

5.9 Intrinsically disordered regions are not exclusive to PduMCP

A few recent reports highlight the importance of disordered regions/proteins in the assembly of carboxysomes in cyanobacteria. The disordered protein CsoS2 is shown to drive the phase separation of RuBisCo of α-carboxysome [66]. Similarly, the disordered SSUL domain in CcmM apparently binds to RuBisCo driving phase separation during ß-carboxysome assembly [65, 119]. Our findings on the distribution and roles of disordered regions in propane diol metabolosome suggests that these regions may also play a role in other metabolosmes. We looked for the presence of intrinsically disordered regions among the ethanolamine utilization microcompartment (EUT) proteins. We obtained the amino acids sequence of all the EUT proteins from NCBI protein database and queried them using the above-mentioned webservers. Four EUT proteins were predicted to have more than 30% of their residues falling in the intrinsically disordered regions **(Table S2)**, shell protein EutK, enzymes EutT and EutC, and a protein with unknown function EutQ. EutK has an extended C-terminal region similar to PduK, however, unlike PduK, the C-terminal domain of EutK does not have extensive disordered region. The C-terminal extension of EutK has helix-turn-helix motif which is usually seen in DNA/RNA binding proteins [124]. IUpred and ANCHOR webservers identified a binding region (Glu111-Ser119). This binding region is rich in hydrophobic residues (ALLALLA) and is preceded by a stretch of intrinsically disordered region (A93-P107). **(Fig. 5.11a)**. The crystal structure of the C-terminal domain of EutK from *Escherichia coli* (strain K12) is available (PDB ID: 3I71) [38]. The structure clearly shows that the hydrophobic residues LLALL indeed lie in the α-helical region **(Fig. 5.11b)**. To further validate our predictions, we modelled the C-terminal domain of EutK (from *Salmonella enterica*) using webserver Galaxy server. The model of EutK C-terminus shows a short-disordered stretch followed by helix-turn-helix motif containing the 'ALLALLA' binding region **(Fig. 5.11c)**, supporting ANCHOR predictions. The function of helix-turn-helix motif in the C-terminal region of EutK is not understood. A thorough examination is required to determine whether the helix-turn-helix pattern is involved in binding DNA. In fact, the arginine-rich motif (RKRSSRYR) was found to be the most likely

DNA binding stretch by our DNA binding prediction using the DRNApred webserver [125]. Two sequential ß-strands make up this motif, which is located beyond the helix-turn-helix motif **(Fig. 5.11d)**. Although it's possible that the positively charged residues may bind to the DNA's negatively charged backbone, the hydrophobic patch suggests that EutK would also bind to other proteins. Moreover, the preceding disordered region would give this binding area the necessary flexibility to engage with partner molecules. However, the binding partners of EutK is not known in the literature.

a) MINALGLLEVDGMVAAVDAADAMLKAA NVRLLSHQVLDPGRLTLVVEGDLAACR AALDAGSAAAQHTGRVISRKEIGRPEE DTQWLIGGFAR**ATTPTEKAPQVPATP**E FA**EALLALLAS**VRQGMTAGEVAAHFGW PLEQARNVLEQLFSDGALRKRSSRYRI KN

Predicted DNA binding "RKRSSRYR" motif
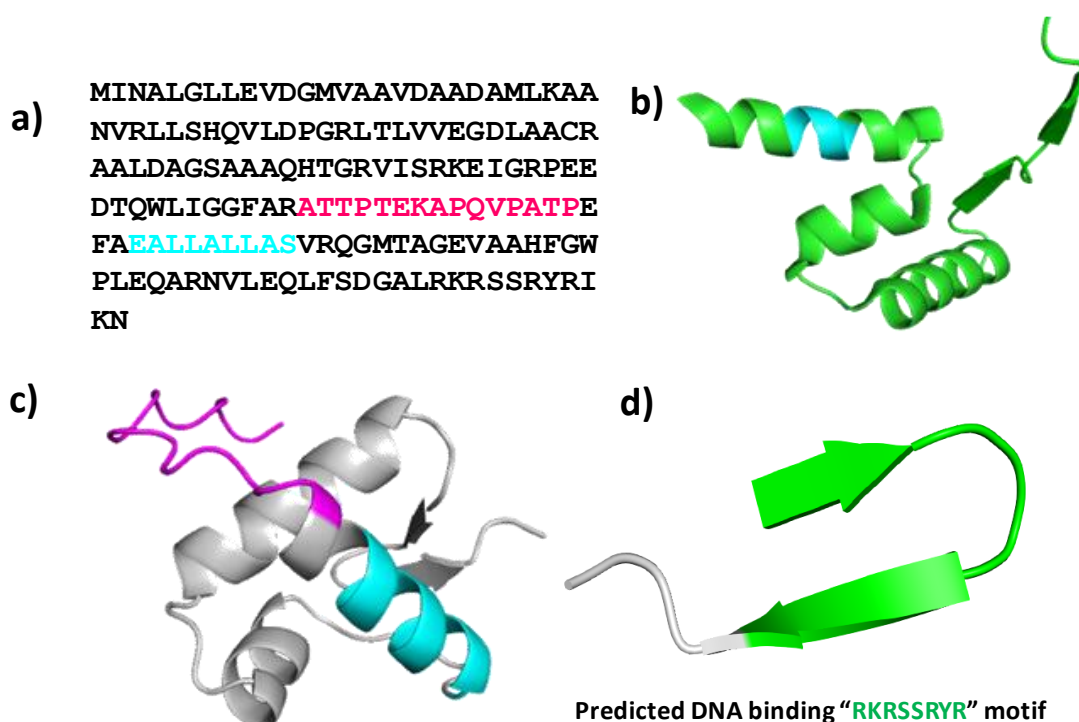
Fig. 5.11: a) Amino acid sequence of EutK (*Salmonella enterica*) showing predicted disordered region (in magenta) and binding region (in cyan) predicted by IUpred and ANCHOR webserver, b) Crystal structure of C-terminal domain of EutK from *E. coli* (PDB ID: 3I71), with residues LLALL highlighted in cyan. c) Model of C-terminal region of EutK from S*almonella enterica* (modeling performed Galaxy TBM server, showing disordered region (in magenta) and binding region (in cyan), d) DNA binding segment in the C-terminal of EutK, predicted using DRNApred webserver.

Table 5.4: Percentage of disordered regions in shell protein and enzymes of EutMCP			
Eut-MCP Proteins	Function	No. of Residues	% disorder PONDR-VSL2
EutS	Shell protein	111	9.91
EutP	Unknown function	159	16.98
EutQ	Unknown function	229	41.92
EutT	Corrinoid adenosyltransferase	267	41.95
EutD	Phosphotransacylase	338	18.34
EutM	Shell protein	96	11.46
EutN	Shell protein	95	17.89
EutE	Alcohol dehydrogenase	467	13.70
EutJ	Chaperon	279	11.83
EutG	Alcohol dehydrogenase	395	8.86
EutH	Permease	408	8.58
EutA	Ethanolamine ammonia lyase reactivase	467	5.14
EutB	Ethanolamine ammonia lyase	453	5.74
EutC	Ethanolamine ammonia lyase	298	33.56
EutL	Shell protein	219	9.59
EutK	Shell protein	164	34.76
EutR	Transcription activator	350	14.29

5.10 Conclusion

The present work predicts the distribution of IDRs in the shell proteins and enzymes of PduMCP. The presence of disordered regions in PduMCP constituent proteins suggests that they may play a role in mediating shell enzyme interaction, allowing shell proteins to be more

flexible, and optimizing their self-assembly behaviour. While the disordered terminal extensions in hexametric (PduA/PduK) or trimeric shell proteins (PduBB') may allow binding with partner proteins or metabolites, the presence of short disordered segments in pentameric vertex shell protein (PduN) may allow structural flexibility to the shell protein. The structural flexibility of vertex protein would allow it to better associate with multiple shell protein partners at the vertices. In the literature, the role of N-terminal extensions in shell-enzyme interaction has been suggested. For example, the N-terminal region of PduBB' has been shown to facilitate better encapsulation of enzymes within PduMCP [47]. Similarly, the N-terminal extension of PduD has been suggested to be crucial for PduCDE encapsulation [79]. Our simulation results show that the disorder in N-terminal extensions PduB and PduD would lead to a flexible association between the two proteins. A flexible association between shell protein and enzyme would make the enzymatic core more dynamic and less rigid. The absence of disordered N-terminal region in PduD weakens its association with the N-terminal extension of PduB, supporting our previous report.

Earlier we have shown that the N-terminal extension of PduB regulates its solubility by counteracting PduB's high self-associating tendency, giving PduBB' combination a well-balanced self-assembly property [67]. This idea is supported by the present finding using limited proteolysis, which shows that the longer version of the shell-protein (PduB) is more soluble and accessible to proteolytic cleavage than the N-terminally truncated PduB'. The observation that the flexible N-terminal extension with disordered segment provides solubility to PduBB' is in agreement with the concept of 'entropic bristle' [126, 127]. Flexible disordered extensions in proteins have the potential to increase the solubility of proteins by excluding large volume around them and by displacing large molecules like neighbouring protein molecules in solution. This concept has been used in the literarure to solubilize recalcitrant proteins like CTLA4 [126]. As suggested by CABS Flex simulation data and limited proteolysis study, the N-terminal region of the shell protein has high accessibility to solvent and exhibits high flexibility. The N-terminal extension of PduBB' would move randomly about its point of attachment and expel away nearby PduBB' molecules. Therefore, the N-terminal region of PduB (longer component) may impede the self-association of PduBB' by excluding large volumes around the shell protein and exerting steric hindrance, countering the high self-associating propensity of PduB' (truncated component). Beside PduBB', the presence of disordered terminal regions in other shell proteins like PduK or PduA, point towards their apparent role in binding enzyme or

metabolite partners. Although the C-terminal of PduA has been suggested to bind PduP, the exact function of PduK C-terminal remains obscure since PduP binding with PduK occurs even in the absence of C-terminal of PduK [120].

We find that disordered regions are not only found in PduMCP but also in EutMCPs, indicating their possible widespread presence in metabolosomes. Recent reports on carboxysomes and pyrenoids have highlighted the role of disordered regions in the phase separation of RuBisCo enzymes during their self-assembly. For example, RuBisCo phase separation in pyrenoids has been shown be mediated by disordered protein EPYC1 [128]. Similarly, in case of ß-carboxysome, RuBisCo binding with the disordered RuBisCo binding domain of CcmM shell protein drives phase separation [65]. Notably, the disordered regions in PduMCP proteins predicted in our study lack the properties of conventional disordered proteins such as low residual complexity (essential for phase separation). Our results suggest that the disordered regions in shell proteins makes them more flexible and soluble; and may result in a dynamic association with the native enzymes.

Chapter 6

Summary and Conclusion

Compartmentalization or the confinement of molecules and reactions within a well-defined volume played a significant role in the origin of cellular life forms. Sequestering biochemical reactions with compartments helps cells to synergize multiple biochemical pathways and improves cellular functions. Studying subcellular compartments not only broadens our understanding of cellular complexity but also provides a glimpse into the prebiological compartmentalization. Bacterial microcompartments are an interesting paradigm to study the self-assembly of biomacromolecules and their impact on the cellular functions. Diverse role of MCPs have helped wide number of bacterial species to survive contribute to energy recycling in the ecosystem. They withstood the long course of evolution as they helped bacteria to survive better under selection pressure. For example, photosynthetic bacteria produce carboxysomes with encapsulated RuBisCo enzyme. These carboxysomes help in global CO_2 fixation and are credited to lowering the CO_2 concentration in the atmosphere during the initial ages of earth. Throughout the evolutionary process, the catalytic efficiency of RuBisCo remained low and this could have proven to be detrimental to photosynthetic bacteria. However, a few photosynthetic bacteria acquired the mechanism to compartmentalize RuBisCo within carboxysomes. Encapsulation of RuBisCo within a carboxysome helps to retain CO_2 in the vicinity of RuBisCo molecules, improving its catalytic efficacy. The metabolosomes take part in catabolism of substrates and may help certain pathogens to outcompete other bacteria. Whether compartmentalization of enzyme with a metabolosome affect its catalytic efficiency is an interesting question to study and has been detailed out in the previous chapters of this thesis. Besides, understanding the self-assembly of the component proteins and their structural basis is necessary for not only understanding the evolution of complex life forms but also to build a platform for the development of artificial protein based nano-factories for various applications.

In this thesis, I have applied a variety of biophysical approaches to PduMCP, supported by biochemical, molecular, and computational methods. PduMCPs are proteinaceous in nature and assist certain bacteria, such as *Salmonella enterica* LT2, in using 1,2-propanediol as an alternative carbon source in low-energy environments. I have elucidated a few remarkable structural features and functions of a major shell protein PduBB'. Shell protein PduBB' which comprises a major portion of the outer shell of PduMCP is unique in the microcompartment world. It exists as a combination of two proteins: PduB (270 residues) and PduB' (233 residues). The longer component PduB has extra 37 amino acids N-terminal extension. In *Chapter 3*, I have explored the link between molecular confinement and

enzyme stability using enzyme assays and biophysical techniques. I observed that in comparison to bare PduCDE in bulk environment, encapsulated PduCDE remains stable under thermal stress and show higher catalytic activity. Intact PduMCP remains stable up to 60°C. It retains optimum catalytic activity of diol dehydratase PduCDE up to 50°C, while bare PduCDE tends to lose its activity above physiological temperatre of 37°C. The outer shell layer of PduMCP appears to have chaperone like role that protects the encapsulated enzyme from thermal denaturation. The major component of the outer shell PduBB' interacts with the enzyme PduCDE and preserves its activity during heat stress. PduBB" not only prevents aggregation of PduCDE but also enhances its activity under physiological temperature without any thermal stress. Both PduB and PduB' (component proteins of PduBB') fail to show similar protective role and do not prevent aggregation of PduCDE under thermal stress. The significance of PduBB' combination has been investigated by studying the stability and self-assembly of the shell proteins and probing their affinity towards PduCDE. In terms of self-assembly, the longer component PduB displays least self-associating property and lower thermal stability and unfolds around ~60°C. The truncated version of the shell protein PduB' displays high thermal stability and unfolds around ~80°C, however, it remains insoluble in solution phase post purification, indicating high self-associating property. The combination protein PduBB' displays the property of both the individual components. PduBB' unfolds around ~80°C and remains soluble in solution phase. The combination of two proteins in PduBB' makes it thermostable and exhibits a well-balanced self-assembly behavior. Although all the three shell proteins have the potential to interact with the enzyme PduCDE, the combination PduBB' shows comparatively higher affinity. This results in comparatively higher affinity towards PduCDE that may provide necessary scaffolding needed for its thermal stability. This chapter provides an overall idea about how compartmentalization and molecular confinement in metabolosomes would not only improve catalytic function but also enhance the structural stability of the enzyme molecules. Enhanced catalytic efficiency within a metabolosome would contribute towards better survival of bacteria when there is a scarcity of nutrients in the surrounding environment. Also, the protective nature of the shell proteins would provide additional benefit to the bacteria under thermal stress, for example during heat stress in the gut microenvironment.

In *Chapter 4,* while probing the self-assembly behavior of shell protein, I observed that the increase in ionic strength results in self-association of PduBB'. Increasing molecular

crowding using crowding agent in the presence of slats trigger PduBB' to separate out of solution as liquid condensates. Compared to monovalent salts, divalent ions have higher potential to induce liquid-liquid phase separation of PduBB'. The surface charge masking of shell protein, together with kosmotropic effect of metal ions were found to be the possible reasons behind salt mediated self-assembly of shell protein. The individual component of PduBB' (PduB and PduB') show distinct phase separation behavior. The truncated variant PduB' forms solid associates irrespective of the protein concentration. The longer component PduB with N-terminal region displays least phase separation propensity and partial phase separation is seen only when protein concentration is increased upto 1.5 mg/ml and above. PduBB' combination displays the property of both PduB and PduB' forming liquid droplets up to very low protein concentration and transitioning into solid associates at high protein concentration of 2 mg/ml. This observation supports our results from Chapter 3, where we show that the combination protein exhibits the self-assembly property of both its individual component. The co-phase separation of PduBB' with PduCDE improves the catalytic property of the enzyme. While both PduB and PduB' interacts with PduCDE, they fail to influence its activity, further strengthening the idea that the combination PduB and PduB' is necessary for proper scaffolding of PduCDE. The idea that the self-assembly of shell protein and enzyme could be triggered by presence of salts was reflected in our study with purified PduMCP. It was found that the presence of salts was necessary for the stability and optimum functioning of intact PduMCP *in vitro*. Notably, Mg^{2+} ions which displayed excellent ability to induce self-assembly and phase separation, was crucial for the stability and optimum activity of PduMCP. In bacterial cells, monovalent ions like Na^+/K^+ range between 150-400 mM while divalent ions like Mg^{2+} are present in lower concentration (1-10 mM). It is possible that these ions, under crowded/viscous cellular environment could induce the microcompartment proteins to assemble. However, it would be a good idea to explore, if there are specific interactions between metal ions and individual shell proteins of the microcompartment.

In *Chapter 5*, I employed bioinformatics and computational tools to study the distribution and potential role of disordered regions in PduMCP. According to our initial bioinformatics, disordered regions are mostly localized on the terminal regions of the shell proteins. The N-terminal extension of PduBB' has a short helical segment followed by intrinsically disordered region. Similarly, the middle subunit (subunit D) of PduCDE also has a helical portion followed by disordered region. Using Docking and MD simulation studies the importance of

the N-terminal extensions of PduB and PduD was assessed. The results showed that N-terminal extensions of PduB and PduD could mediate a flexible shell protein-enzyme interaction. Deletion of N-terminal region of PduD weakened its binding with the shell protein PduBB'. We propose that the disordered regions provide flexibility to the N-terminal extensions of shell protein and enzyme, facilitating their association mediated by the helical segments. We further attempted to experimentally verify the presence of disordered region in the N-terminal region of PduBB'. Through limited proteolysis experiment, we demonstrate that the N-terminal region of PduBB' is disordered and provides solubility. The disordered N-terminal regions act like entropic bristle that enhances the solubility of protein by preventing nearby protein molecules to self-assemble. Proteolysis of the N-terminal region results in self-association of the shell protein and reduces its susceptibility to proteolytic cleavage. The balanced self-assembly behavior of PduBB' observed in **Chapter 3** and **Chapter 4** may be attributed to the disordered region of the shell protein. The flexible disordered N-terminal of the longer component (PduB) counters the high self-associating property of PduB' (N-terminal truncated component) resulting in a balanced self-assembly behavior. A vital role of the N-terminal of PduB in self-assembly and enzyme binding is both exciting and puzzling at the same time. While the apparent role of PduB N-terminal in cargo encapsulation could be true to *Salmonella* and perhaps in other species like *Klebsiella* or *Yersinia*, the absence of this terminal region in other species is bewildering. It is plausible that terminal regions in other shell proteins could facilitate encapsulation. In a multi-protein system like PduMCP multiple protein might perform similar functions. It is very likely that cargo proteins of PduMCP can adapt depending on the proteins' availability when it comes to binding shell protein partners across species. It would be very helpful to carry out a bioinformatics survey to identify segments or peptides that would perform function similar to the N-terminal of PduB.

The work embodied in this thesis may be summarized into three key findings. First, the major shell protein PduBB' has a chaperon like role towards the signature enzyme PduCDE. The combination of PduB and PduB' is essential for thermal stability and optimized self-assembly behavior of the shell protein. Second, the variation in the environmental stimuli like ionic strength and macromolecular crowding governs the self-assembly of PduBB' and PduCDE. The self-assembly of PduBB" and PduCDE is probably a result of both protein-protein interaction and phase separation. Third, the disordered N-terminal regions of PduBB' and enzyme PduCDE (subunit D) participates by shell-enzyme interaction which is highly flexible and dynamic.

Bibliography

1. Wolkenhauer O, Muir A: **The complexity of cell-biological systems**. In: *Philosophy of complex systems.* Elsevier; 2011: 355-385.

2. Woese CR: **On the evolution of cells**. *Proceedings of the National Academy of Sciences* 2002, **99**(13):8742-8747.

3. Wolf YI, Katsnelson MI, Koonin EV: **Physical foundations of biological complexity**. *Proceedings of the National Academy of Sciences* 2018, **115**(37):E8678-E8687.

4. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: **The compartmentalization of cells**. In: *Molecular Biology of the Cell 4th edition.* Garland Science; 2002.

5. Lebiedzinska M, Szabadkai G, Jones AW, Duszynski J, Wieckowski MR: **Interactions between the endoplasmic reticulum, mitochondria, plasma membrane and other subcellular organelles**. *The international journal of biochemistry & cell biology* 2009, **41**(10):1805-1816.

6. André B: **An overview of membrane transport proteins in Saccharomyces cerevisiae**. *Yeast* 1995, **11**(16):1575-1611.

7. Tan S, Tan HT, Chung MC: **Membrane proteins and membrane proteomics**. *Proteomics* 2008, **8**(19):3924-3932.

8. Gomes E, Shorter J: **The molecular language of membraneless organelles**. *Journal of Biological Chemistry* 2019, **294**(18):7115-7127.

9. Kuechler ER, Budzyńska PM, Bernardini JP, Gsponer J, Mayor T: **Distinct features of stress granule proteins predict localization in membraneless organelles**. *Journal of molecular biology* 2020, **432**(7):2349-2368.

10. Drino A, Schaefer MR: **RNAs, phase separation, and membrane-less organelles: Are post-transcriptional modifications modulating organelle dynamics?** *BioEssays* 2018, **40**(12):1800085.

11. Sehgal PB, Westley J, Lerea KM, DiSenso-Browne S, Etlinger JD: **Biomolecular condensates in cell biology and virology: Phase-separated membraneless organelles (MLOs)**. *Analytical biochemistry* 2020, **597**:113691.

12. Kirst H, Kerfeld CA: **Bacterial microcompartments: catalysis-enhancing metabolic modules for next generation metabolic and biomedical engineering**. *BMC biology* 2019, **17**(1):1-11.

13. Kerfeld CA, Aussignargues C, Zarzycki J, Cai F, Sutter M: **Bacterial microcompartments**. *Nature Reviews Microbiology* 2018, **16**(5):277.

14. Stewart AM, Stewart KL, Yeates TO, Bobik TA: **Advances in the world of bacterial microcompartments**. *Trends in biochemical sciences* 2021, **46**(5):406-416.

15. Kerfeld CA, Heinhorst S, Cannon GC: **Bacterial microcompartments**. *Annual review of microbiology* 2010, **64**(LBNL-3809E).

16. Cheng S, Liu Y, Crowley CS, Yeates TO, Bobik TA: **Bacterial microcompartments: their properties and paradoxes**. *Bioessays* 2008, **30**(11-12):1084-1095.

17. Yeates TO, Thompson MC, Bobik TA: **The protein shells of bacterial microcompartment organelles**. *Current opinion in structural biology* 2011, **21**(2):223-231.

18. Yeates TO, Crowley CS, Tanaka S: **Bacterial microcompartment organelles: protein shell structure and evolution**. *Annual review of biophysics* 2010, **39**:185.

19. Turmo A, Gonzalez-Esquer CR, Kerfeld CA: **Carboxysomes: metabolic modules for CO2 fixation**. *FEMS microbiology letters* 2017, **364**(18).

20. Whitehead L, Long BM, Price GD, Badger MR: **Comparing the in vivo function of α-carboxysomes and β-carboxysomes in two model cyanobacteria**. *Plant Physiology* 2014, **165**(1):398-411.

21. Prentice MB: **Bacterial microcompartments and their role in pathogenicity**. *Current Opinion in Microbiology* 2021, **63**:19-28.

22. Lee MJ, Palmer DJ, Warren MJ: **Biotechnological advances in bacterial microcompartment technology**. *Trends in biotechnology* 2019, **37**(3):325-336.

23. Axen SD, Erbilgin O, Kerfeld CA: **A taxonomy of bacterial microcompartment loci constructed by a novel scoring method**. *PLoS Comput Biol* 2014, **10**(10):e1003898.

24. Cai F, Dou Z, Bernstein SL, Leverenz R, Williams EB, Heinhorst S, Shively J, Cannon GC, Kerfeld CA: **Advances in understanding carboxysome assembly in Prochlorococcus and Synechococcus implicate CsoS2 as a critical component**. *Life* 2015, **5**(2):1141-1171.

25. Schmid MF, Paredes AM, Khant HA, Soyer F, Aldrich HC, Chiu W, Shively JM: **Structure of Halothiobacillus neapolitanus carboxysomes by cryo-electron tomography**. *Journal of molecular biology* 2006, **364**(3):526-535.

26. Iancu CV, Ding HJ, Morris DM, Dias DP, Gonzales AD, Martino A, Jensen GJ: **The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography**. *Journal of molecular biology* 2007, **372**(3):764-773.

27. Rae BD, Long BM, Badger MR, Price GD: **Structural determinants of the outer shell of β-carboxysomes in Synechococcus elongatus PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL**. *PLoS One* 2012, **7**(8):e43871.

28. Shively J, Ball FL, Kline BW: **Electron microscopy of the carboxysomes (polyhedral bodies) of Thiobacillus neapolitanus**. *Journal of Bacteriology* 1973, **116**(3):1405-1411.

29. Kaneko Y, Danev R, Nagayama K, Nakamoto H: **Intact carboxysomes in a cyanobacterial cell visualized by hilbert differential contrast transmission electron microscopy**. *Journal of bacteriology* 2006, **188**(2):805-808.

30. Rae BD, Long BM, Badger MR, Price GD: **Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO2 fixation in cyanobacteria and some proteobacteria**. *Microbiology and Molecular Biology Reviews* 2013, **77**(3):357-379.

31. Faulkner M, Rodriguez-Ramos J, Dykes GF, Owen SV, Casella S, Simpson DM, Beynon RJ, Liu L-N: **Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes**. *Nanoscale* 2017, **9**(30):10662-10673.

32. Rae BD, Long BM, Whitehead LF, Förster B, Badger MR, Price GD: **Cyanobacterial carboxysomes: microcompartments that facilitate CO2 fixation**. *Journal of Molecular Microbiology and Biotechnology* 2013, **23**(4-5):300-307.

33. Iancu CV, Morris DM, Dou Z, Heinhorst S, Cannon GC, Jensen GJ: **Organization, structure, and assembly of α-carboxysomes determined by electron cryotomography of intact cells**. *Journal of molecular biology* 2010, **396**(1):105-117.

34. Drews G, Niklowitz W: **Cytology of Cyanophycea. II. Centroplasm and granular inclusions of Phormidium uncinatum**. *Archiv Fur Mikrobiologie* 1956, **24**(2):147-162.

35. Shively J, Ball F, Brown D, Saunders R: **Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus**. *Science* 1973, **182**(4112):584-586.

36. Chen P, Andersson DI, Roth JR: **The control region of the pdu/cob regulon in Salmonella typhimurium**. *Journal of bacteriology* 1994, **176**(17):5474-5482.

37. Stojiljkovic I, Bäumler AJ, Heffron F: **Ethanolamine utilization in Salmonella typhimurium: nucleotide sequence, protein expression, and mutational analysis of the cchA cchB eutE eutJ eutG eutH gene cluster**. *Journal of bacteriology* 1995, **177**(5):1357-1366.

38. Tanaka S, Sawaya MR, Yeates TO: **Structure and mechanisms of a protein-based organelle in Escherichia coli**. *Science* 2010, **327**(5961):81-84.

39. Pang A, Liang M, Prentice MB, Pickersgill RW: **Substrate channels revealed in the trimeric Lactobacillus reuteri bacterial microcompartment shell protein PduB**. *Acta Crystallographica Section D: Biological Crystallography* 2012, **68**(12):1642-1652.

40. Jorda J, Lopez D, Wheatley NM, Yeates TO: **Using comparative genomics to uncover new kinds of protein-based metabolic organelles in bacteria**. *Protein Science* 2013, **22**(2):179-195.

41. Lundin AP, Stewart KL, Stewart AM, Herring TI, Chowdhury C, Bobik TA: **Genetic characterization of a glycyl radical microcompartment used for 1, 2-propanediol fermentation by uropathogenic Escherichia coli CFT073**. *Journal of bacteriology* 2020, **202**(9):e00017-00020.

42. Seedorf H, Fricke WF, Veith B, Brüggemann H, Liesegang H, Strittmatter A, Miethke M, Buckel W, Hinderberger J, Li F: **The genome of Clostridium kluyveri, a strict**

anaerobe with unique metabolic features. *Proceedings of the National Academy of Sciences* 2008, **105**(6):2128-2133.

43.   Herring TI, Harris TN, Chowdhury C, Mohanty SK, Bobik TA: **A bacterial microcompartment is used for choline fermentation by Escherichia coli 536**. *Journal of bacteriology* 2018, **200**(10):e00764-00717.

44.   Park J, Chun S, Bobik TA, Houk KN, Yeates TO: **Molecular dynamics simulations of selective metabolite transport across the propanediol bacterial microcompartment shell**. *The Journal of Physical Chemistry B* 2017, **121**(34):8149-8154.

45.   Chowdhury C, Sinha S, Chun S, Yeates TO, Bobik TA: **Diverse bacterial microcompartment organelles**. *Microbiology and Molecular Biology Reviews* 2014, **78**(3):438-468.

46.   Cheng S, Sinha S, Fan C, Liu Y, Bobik TA: **Genetic analysis of the protein shell of the microcompartments involved in coenzyme B12-dependent 1, 2-propanediol degradation by Salmonella**. *Journal of bacteriology* 2011, **193**(6):1385-1392.

47.   Lehman BP, Chowdhury C, Bobik TA: **The N terminus of the PduB protein binds the protein shell of the Pdu microcompartment to its enzymatic core**. *Journal of bacteriology* 2017, **199**(8):e00785-00716.

48.   Long BM, Tucker L, Badger MR, Price GD: **Functional cyanobacterial β-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein**. *Plant Physiology* 2010, **153**(1):285-293.

49.   Kinney JN, Salmeen A, Cai F, Kerfeld CA: **Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly**. *Journal of Biological Chemistry* 2012, **287**(21):17729-17736.

50.   Kumar G, Sinha S: **Biophysical approaches to understand and re-purpose bacterial microcompartments**. *Current Opinion in Microbiology* 2021, **63**:43-51.

51.   Kennedy NW, Hershewe JM, Nichols TM, Roth EW, Wilke CD, Mills CE, Jewett MC, Tullman-Ercek D: **Apparent size and morphology of bacterial microcompartments varies with technique**. *PloS one* 2020, **15**(3):e0226395.

52. Sutter M, Greber B, Aussignargues C, Kerfeld CA: **Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell**. *Science* 2017, **356**(6344):1293-1297.

53. Kim EY, Tullman-Ercek D: **A rapid flow cytometry assay for the relative quantification of protein encapsulation into bacterial microcompartments**. *Biotechnology Journal* 2014, **9**(3):348-354.

54. Yang M, Simpson DM, Wenner N, Brownridge P, Harman VM, Hinton JC, Beynon RJ, Liu L-N: **Decoding the stoichiometric composition and organisation of bacterial metabolosomes**. *Nature communications* 2020, **11**(1):1-11.

55. Sinha S, Cheng S, Sung YW, McNamara DE, Sawaya MR, Yeates TO, Bobik TA: **Alanine scanning mutagenesis identifies an asparagine–arginine–lysine triad essential to assembly of the shell of the Pdu microcompartment**. *Journal of molecular biology* 2014, **426**(12):2328-2345.

56. Pang A, Frank S, Brown I, Warren MJ, Pickersgill RW: **Structural insights into higher order assembly and function of the bacterial microcompartment protein PduA**. *Journal of Biological Chemistry* 2014, **289**(32):22377-22384.

57. Uddin I, Frank S, Warren MJ, Pickersgill RW: **A generic self-assembly process in microcompartments and synthetic protein nanotubes**. *Small* 2018, **14**(19):1704020.

58. Noël CR, Cai F, Kerfeld CA: **Purification and characterization of protein nanotubes assembled from a single bacterial microcompartment shell subunit**. *Advanced Materials Interfaces* 2016, **3**(1):1500295.

59. Sutter M, Faulkner M, Aussignargues C, Paasch BC, Barrett S, Kerfeld CA, Liu L-N: **Visualization of bacterial microcompartment facet assembly using high-speed atomic force microscopy**. *Nano letters* 2016, **16**(3):1590-1595.

60. Faulkner M, Zhao L-S, Barrett S, Liu L-N: **Self-assembly stability and variability of bacterial microcompartment shell proteins in response to the environmental change**. *Nanoscale research letters* 2019, **14**(1):1-8.

61. Bari NK, Hazra JP, Kumar G, Kaur S, Sinha S: **Probe into a multi-protein prokaryotic organelle using thermal scanning assay reveals distinct properties of**

the core and the shell. *Biochimica et Biophysica Acta (BBA)-General Subjects* 2020, **1864**(10):129680.

62. Mayer MJ, Juodeikis R, Brown IR, Frank S, Palmer DJ, Deery E, Beal DM, Xue W-F, Warren MJ: **Effect of bio-engineering on size, shape, composition and rigidity of bacterial microcompartments**. *Scientific reports* 2016, **6**(1):1-11.

63. Cai F, Sutter M, Cameron JC, Stanley DN, Kinney JN, Kerfeld CA: **The structure of CcmP, a tandem bacterial microcompartment domain protein from the β-carboxysome, forms a subcompartment within a microcompartment**. *Journal of Biological Chemistry* 2013, **288**(22):16055-16063.

64. Cameron JC, Wilson SC, Bernstein SL, Kerfeld CA: **Biogenesis of a bacterial organelle: the carboxysome assembly pathway**. *Cell* 2013, **155**(5):1131-1140.

65. Wang H, Yan X, Aigner H, Bracher A, Nguyen ND, Hee WY, Long B, Price GD, Hartl F, Hayer-Hartl M: **Rubisco condensate formation by CcmM in β-carboxysome biogenesis**. *Nature* 2019, **566**(7742):131-135.

66. Oltrogge LM, Chaijarasphong T, Chen AW, Bolin ER, Marqusee S, Savage DF: **Multivalent interactions between CsoS2 and Rubisco mediate α-carboxysome formation**. *Nature structural & molecular biology* 2020, **27**(3):281-287.

67. Kumar G, Bari NK, Hazra JP, Sinha S: **A Major Shell Protein of 1, 2-Propanediol Utilization Microcompartment Conserves the Activity of Its Signature Enzyme at Higher Temperatures**. *ChemBioChem* 2022:e202100694.

68. Indirapriyadharshini V, Ramamurthy P: **Fluorescence anisotropy of acridinedione dyes in glycerol: Prolate model of ellipsoid**. *Journal of Chemical Sciences* 2007, **119**(2):161-168.

69. Shin W-H, Lee GR, Heo L, Lee H, Seok C: **Prediction of protein structure and interaction by GALAXY protein modeling programs**. *Bio Design* 2014, **2**(1):1-11.

70. Seok C, Baek M, Steinegger M, Park H, Lee GR, Won J: **Accurate protein structure prediction: what comes next?** *Biodesign* 2021, **9**:47-50.

71. Kurcinski M, Jamroz M, Blaszczyk M, Kolinski A, Kmiecik S: **CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site**. *Nucleic acids research* 2015, **43**(W1):W419-W424.

72. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A: **PRODIGY: a web server for predicting the binding affinity of protein–protein complexes**. *Bioinformatics* 2016, **32**(23):3676-3678.

73. Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN: **PONDR-FIT: a meta-predictor of intrinsically disordered amino acids**. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 2010, **1804**(4):996-1010.

74. Mészáros B, Erdős G, Dosztányi Z: **IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding**. *Nucleic acids research* 2018, **46**(W1):W329-W337.

75. Dosztányi Z, Mészáros B, Simon I: **ANCHOR: web server for predicting protein binding regions in disordered proteins**. *Bioinformatics* 2009, **25**(20):2745-2746.

76. Disfani FM, Hsu W-L, Mizianty MJ, Oldfield CJ, Xue B, Dunker AK, Uversky VN, Kurgan L: **MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins**. *Bioinformatics* 2012, **28**(12):i75-i83.

77. Kuriata A, Gierut AM, Oleniecki T, Ciemny MP, Kolinski A, Kurcinski M, Kmiecik S: **CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures**. *Nucleic acids research* 2018, **46**(W1):W338-W343.

78. Kurcinski M, Badaczewska-Dawid A, Kolinski M, Kolinski A, Kmiecik S: **Flexible docking of peptides to proteins using CABS-dock**. *Protein Science* 2020, **29**(1):211-222.

79. Fan C, Bobik TA: **The N-terminal region of the medium subunit (PduD) packages adenosylcobalamin-dependent diol dehydratase (PduCDE) into the Pdu microcompartment**. *Journal of bacteriology* 2011, **193**(20):5623-5628.

80. Monnard P-A, Walde P: **Current ideas about prebiological compartmentalization**. *Life* 2015, **5**(2):1239-1263.

81. Plegaria JS, Kerfeld CA: **Engineering nanoreactors using bacterial microcompartment architectures**. *Current opinion in biotechnology* 2018, **51**:1-7.

82. Bari NK, Kumar G, Bhatt A, Hazra JP, Garg A, Ali ME, Sinha S: **Nanoparticle fabrication on bacterial microcompartment surface for the development of hybrid enzyme-inorganic catalyst**. *ACS Catalysis* 2018, **8**(9):7742-7748.

83. Sampson EM, Bobik TA: **Microcompartments for B12-dependent 1, 2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate**. *Journal of bacteriology* 2008, **190**(8):2966-2971.

84. Havemann GD, Bobik TA: **Protein content of polyhedral organelles involved in coenzyme B12-dependent degradation of 1, 2-propanediol in Salmonella enterica serovar Typhimurium LT2**. *Journal of bacteriology* 2003, **185**(17):5086-5095.

85. Tocheva EI, Matson EG, Cheng SN, Chen WG, Leadbetter JR, Jensen GJ: **Structure and expression of propanediol utilization microcompartments in Acetonema longum**. *Journal of bacteriology* 2014, **196**(9):1651-1658.

86. Chowdhury C, Sinha S, Chun S, Yeates TO, Bobik TA: **Diverse bacterial microcompartment organelles**. *Microbiol Mol Biol Rev* 2014, **78**(3):438-468.

87. Bobik TA, Havemann GD, Busch RJ, Williams DS, Aldrich HC: **The Propanediol Utilization (pdu) Operon ofSalmonella enterica Serovar Typhimurium LT2 Includes Genes Necessary for Formation of Polyhedral Organelles Involved in Coenzyme B12-Dependent 1, 2-Propanediol Degradation**. *Journal of bacteriology* 1999, **181**(19):5967-5975.

88. Kim EY, Slininger MF, Tullman-Ercek D: **The effects of time, temperature, and pH on the stability of PDU bacterial microcompartments**. *Protein Science* 2014, **23**(10):1434-1441.

89. Choudhary D, Kumar A, Magliery TJ, Sotomayor M: **Using thermal scanning assays to test protein-protein interactions of inner-ear cadherins**. *PLoS One* 2017, **12**(12):e0189546.

90. Jakobson CM, Tullman-Ercek D, Slininger MF, Mangan NM: **A systems-level model reveals that 1, 2-Propanediol utilization microcompartments enhance pathway flux through intermediate sequestration**. *PLoS computational biology* 2017, **13**(5):e1005525.

91. Eggers DK, Valentine JS: **Molecular confinement influences protein structure and enhances thermal protein stability**. *Protein Science* 2001, **10**(2):250-261.

92. Bolivar JM, Wilson L, Ferrarotti SA, Fernandez-Lafuente R, Guisan JM, Mateo C: **Stabilization of a formate dehydrogenase by covalent immobilization on highly activated glyoxyl-agarose supports**. *Biomacromolecules* 2006, **7**(3):669-673.

93. Bayramoğlu G, Yilmaz M, Arica MY: **Immobilization of a thermostable α-amylase onto reactive membranes: kinetics characterization and application to continuous starch hydrolysis**. *Food Chemistry* 2004, **84**(4):591-599.

94. Ravcheev DA, Moussu L, Smajic S, Thiele I: **Comparative genomic analysis reveals novel microcompartment-associated metabolic pathways in the human gut microbiome**. *Frontiers in genetics* 2019, **10**:636.

95. Chowdhury C, Chun S, Sawaya MR, Yeates TO, Bobik TA: **The function of the PduJ microcompartment shell protein is determined by the genomic position of its encoding gene**. *Molecular microbiology* 2016, **101**(5):770-783.

96. Crowley CS, Sawaya MR, Bobik TA, Yeates TO: **Structure of the PduU shell protein from the Pdu microcompartment of Salmonella**. *Structure* 2008, **16**(9):1324-1332.

97. Bolognesi B, Gotor NL, Dhar R, Cirillo D, Baldrighi M, Tartaglia GG, Lehner B: **A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression**. *Cell reports* 2016, **16**(1):222-231.

98. Van Den Berg J, Boersma AJ, Poolman B: **Microorganisms maintain crowding homeostasis**. *Nature Reviews Microbiology* 2017, **15**(5):309-318.

99. Feig M, Yu I, Wang P-h, Nawrocki G, Sugita Y: **Crowding in cellular environments at an atomistic level from computer simulations**. *The Journal of Physical Chemistry B* 2017, **121**(34):8009-8025.

100. André AA, Spruijt E: **Liquid–liquid phase separation in crowded environments**. *International Journal of Molecular Sciences* 2020, **21**(16):5908.

101. Roque A, Ponte I, Suau P: **Macromolecular crowding induces a molten globule state in the C-terminal domain of histone H1**. *Biophysical journal* 2007, **93**(6):2170-2177.

102. Park S, Barnes R, Lin Y, Jeon B-j, Najafi S, Delaney KT, Fredrickson GH, Shea J-E, Hwang DS, Han S: **Dehydration entropy drives liquid-liquid phase separation by molecular crowding**. *Communications Chemistry* 2020, **3**(1):1-12.

103. Kumar G, Bari NK, Hazra JP, Sinha S: **A major shell protein of 1, 2-propanediol utilization microcompartment conserves the activity of its signature enzyme at higher temperatures**. *ChemBioChem*.

104. 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**. *Nature communications* 2017, **8**(1):1-13.

105. Ray S, Singh N, Kumar R, Patel K, Pandey S, Datta D, Mahato J, Panigrahi R, Navalkar A, Mehra S: **α-Synuclein aggregation nucleates through liquid–liquid phase separation**. *Nature chemistry* 2020, **12**(8):705-716.

106. Kanaan NM, Hamel C, Grabinski T, Combs B: **Liquid-liquid phase separation induces pathogenic tau conformations in vitro**. *Nature communications* 2020, **11**(1):1-16.

107. Wegmann S, Eftekharzadeh B, Tepper K, Zoltowska KM, Bennett RE, Dujardin S, Laskowski PR, MacKenzie D, Kamath T, Commins C: **Tau protein liquid–liquid phase separation can initiate tau aggregation**. *The EMBO journal* 2018, **37**(7):e98049.

108. Schuster BS, Reed EH, Parthasarathy R, Jahnke CN, Caldwell RM, Bermudez JG, Ramage H, Good MC, Hammer DA: **Controllable protein phase separation and modular recruitment to form responsive membraneless organelles**. *Nature communications* 2018, **9**(1):1-12.

109. Bracha D, Walls MT, Brangwynne CP: **Probing and engineering liquid-phase organelles**. *Nature biotechnology* 2019, **37**(12):1435-1445.

110. He S, Chou H-T, Matthies D, Wunder T, Meyer MT, Atkinson N, Martinez-Sanchez A, Jeffrey PD, Port SA, Patena W: **The structural basis of Rubisco phase separation in the pyrenoid**. *Nature plants* 2020, **6**(12):1480-1490.

111. Tange H, Ishibashi D, Nakagaki T, Taguchi Y, Kamatari YO, Ozawa H, Nishida N: **Liquid–liquid phase separation of full-length prion protein initiates conformational conversion in vitro**. *Journal of Biological Chemistry* 2021, **296**.

112. Guilhas B, Walter J-C, Rech J, David G, Walliser NO, Palmeri J, Mathieu-Demaziere C, Parmeggiani A, Bouet J-Y, Le Gall A: **ATP-driven separation of liquid phase condensates in bacteria**. *Molecular Cell* 2020, **79**(2):293-303. e294.

113. Babu MM: **The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease**. *Biochemical Society Transactions* 2016, **44**(5):1185-1200.

114. Van Der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT: **Classification of intrinsically disordered regions and proteins**. *Chemical reviews* 2014, **114**(13):6589-6631.

115. Latonen L: **Phase-to-phase with nucleoli–stress responses, protein aggregation and novel roles of RNA**. *Frontiers in Cellular Neuroscience* 2019, **13**:151.

116. Protter DS, Rao BS, Van Treeck B, Lin Y, Mizoue L, Rosen MK, Parker R: **Intrinsically disordered regions can contribute promiscuous interactions to RNP granule assembly**. *Cell reports* 2018, **22**(6):1401-1412.

117. Luo Y, Na Z, Slavoff SA: **P-bodies: composition, properties, and functions**. *Biochemistry* 2018, **57**(17):2424-2431.

118. Harmon TS, Holehouse AS, Rosen MK, Pappu RV: **Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins**. *elife* 2017, **6**:e30294.

119. Zang K, Wang H, Hartl FU, Hayer-Hartl M: **Scaffolding protein CcmM directs multiprotein phase separation in β-carboxysome biogenesis**. *Nature Structural & Molecular Biology* 2021, **28**(11):909-922.

120. Lawrence AD, Frank S, Newnham S, Lee MJ, Brown IR, Xue W-F, Rowe ML, Mulvihill DP, Prentice MB, Howard MJ: **Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor**. *ACS synthetic biology* 2014, **3**(7):454-465.

121. Fan C, Cheng S, Liu Y, Escobar CM, Crowley CS, Jefferson RE, Yeates TO, Bobik TA: **Short N-terminal sequences package proteins into bacterial microcompartments**. *Proceedings of the National Academy of Sciences* 2010, **107**(16):7509-7514.

122. Nordyke CT, Ahmed YM, Puterbaugh RZ, Bowman GR, Varga K: **Intrinsically disordered bacterial polar organizing protein Z, PopZ, interacts with protein binding partners through an N-terminal Molecular Recognition Feature**. *Journal of molecular biology* 2020, **432**(23):6092-6107.

123. Weng G, Wang E, Wang Z, Liu H, Zhu F, Li D, Hou T: **HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA**. *Nucleic acids research* 2019, **47**(W1):W322-W330.

124. Corbella M, Liao Q, Moreira C, Parracino A, Kasson PM, Kamerlin SCL: **The N-terminal helix-turn-helix motif of transcription factors MarA and Rob drives DNA recognition**. *The Journal of Physical Chemistry B* 2021, **125**(25):6791-6806.

125. Yan J, Kurgan L: **DRNApred, fast sequence-based method that accurately predicts and discriminates DNA-and RNA-binding residues**. *Nucleic acids research* 2017, **45**(10):e84-e84.

126. Santner AA, Croy CH, Vasanwala FH, Uversky VN, Van Y-YJ, Dunker AK: **Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression**. *Biochemistry* 2012, **51**(37):7250-7262.

127. Uversky VN: **The most important thing is the tail: multitudinous functionalities of intrinsically disordered protein termini**. *FEBS letters* 2013, **587**(13):1891-1901.

128. Mackinder LC, Meyer MT, Mettler-Altmann T, Chen VK, Mitchell MC, Caspari O, Rosenzweig ESF, Pallesen L, Reeves G, Itakura A: **A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle**. *Proceedings of the National Academy of Sciences* 2016, **113**(21):5958-5963.

Appendix

Permissions from journals for reuse of content in thesis

Permission to reuse content from an article published by Portland Press:

- If the content that you are seeking to re-use is in a Portland Press article that is published open access under a CC BY licence NO permissions are required, although you must cite the published article and credit the authors when you re-use it (or part of it).

- If the article you are seeking to re-use is published open access under any other type of licence (e.g. CC BY NC-ND) or a Portland Press license to publish then please complete a re-use permission-request form via copyright.com.

- To find out what licence the article is published under look for the copyright line on the published article, which can be found underneath the abstract or full text, depending on what view you are seeing for the article. Where there is no creative commons license attached, please complete a re-use permission request form via copyright.com.

- **FOR AUTHORS:** if you are a named author on the article you wish to re-use then you will not need to seek any permissions except for re-use of non-open access papers that involves commercial re-selling or bulk distribution. For the latter, please visit copyright.com.

☑ Licensed Content

Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	ChemBioChem
Licensed Content Title	A Major Shell Protein of 1,2-Propanediol Utilization Microcompartment Conserves the Activity of Its Signature Enzyme at Higher Temperatures
Licensed Content Author	Gaurav Kumar, Naimat K. Bari, Jagadish P. Hazra, et al
Licensed Content Date	Mar 14, 2022
Licensed Content Volume	23
Licensed Content Issue	9
Licensed Content Pages	8

📄 Order Details

Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No

📄 About Your Work

Title	Functional Attributes and Self-Assembly Behavior of a Shell Protein from the 1,2-Propanediol Utilization Prokaryotic Metabolosome
Institution name	Institute of Nanoscience and Technology (INST) Mohali
Expected presentation date	Feb 2023

📁 Additional Data

Order reference number	Amajorshellgk

📍 Requestor Location

Requestor Location	Mr. Gaurav Kumar Institute of Nanoscience and Technology Phase 10, sector 64 Mohali, Please Select 160062 India Attn: Mr. Gaurav Kumar

📱 Tax Details

Publisher Tax ID	EU826007151

💳 Billing Information

Billing Type	Invoice
Billing address	Mr. Gaurav Kumar Institute of Nanoscience and Technology Phase 10, sector 64 Mohali, India 160062 Attn: Mr. Gaurav Kumar

$ Price

Total	0.00 USD

Biophysical approaches to understand and re-purpose bacterial microcompartments

Author: Gaurav Kumar, Sharmistha Sinha

Publication: Current Opinion in Microbiology

Publisher: Elsevier

Date: October 2021

Journal Author Rights

BACK

CLOSE WINDOW

Disordered regions endow structural flexibility to shell proteins and function towards shell–enzyme interactions in 1,2-propanediol utilization microcompartment

Author: Gaurav Kumar, , Jagadish Prasad Hazra, et al

Publication: Journal of Biomolecular Structure and Dynamics

Publisher: Taylor & Francis

Date: Nov 1, 2022

Thesis/Dissertation Reuse Request

BACK

CLOSE

Publications included in thesis

Kumar G, Bari NK, Hazra JP, Sinha S. A Major Shell Protein of 1, 2-Propanediol Utilization Microcompartment Conserves the Activity of Its Signature Enzyme at Higher Temperatures. ChemBioChem. 2022 May 4: e202100694.

Kumar G, Sinha S. Biophysical approaches to understand and re-purpose bacterial microcompartments. Current Opinion in Microbiology. 2021 Oct 1; 63:43-51.

Kumar G, Hazra JP, Sinha S. Disordered regions endow structural flexibility to shell proteins and function towards shell–enzyme interactions in 1, 2-propanediol utilization microcompartment. Journal of Biomolecular Structure and Dynamics. 2022 Oct 23:1-1.

Kumar G, Sinha S. Self-assembly of Shell Protein and Native Enzyme in a Crowded Environment Leads to Catalytically Active Phase Condensates. Biochemical Journal. 2023 Jan 23:BCJ20220551.