Deciphering the Implications of Modular Rearrangements and Circular Permutations in 1,2 -Propanediol Bacterial Microcompartment Domain Proteins

Simerpreet Kaur

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.

March 2024

Dedicated to my

Parents

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 the contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgment of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

SIMERPREET KAUR

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

Acknowledgment

The journey toward the completion of my Ph.D. thesis has been an invaluable experience for me that has helped me gain a more realistic perspective on life. Along with many challenges all the way through, I was fortunate enough to have some people around me who have helped me and supported me and made this journey smoother. Towards the submission of this thesis, I would like to take the opportunity to deliver my sincere gratitude and thanks to all the people who made it possible and without whom this journey would have been much more difficult.

First and foremost, I would like to express my sincere gratitude to my supervisor **Dr. Sharmistha Sinha** for giving me the invaluable opportunity to be a part of her research group and having confidence in my abilities. Her unwavering support, especially during the initial years, has been a great source of motivation for me. She has always encouraged me to be an independent researcher and allowed me to explore different avenues in my work. Dr. Sinha's constructive criticism has helped us improve our analytical skills, pushing our imaginations beyond limits. I admire her strong personality, optimism, and passion for science. She has been a silent support in many situations, understanding my strengths and weaknesses and guiding me accordingly. I feel extremely fortunate to have had her as my mentor, and I am grateful for all the guidance and encouragement she has provided me with.

I would like to express my gratitude to my thesis committee members, **Dr. Surajit Karmakar** and **Dr. Ehesan Ali,** for their valuable time, interest in my work, and encouragement. I would also like to extend my gratitude to **Dr. Ramya T.N.C**., an external reviewing committee member from CSIR-IMTECH, Chandigarh, for reviewing my work and providing me with insightful suggestions. Their critical reviews and feedback helped me to improve the overall quality of my thesis.

I express my sincere gratitude to **Dr. Sabyasachi Rakshit**, IISER Mohali for his constant support in allowing me to work in his lab during my PhD. I appreciate his kindness and his positive outlook on life which is always encouraging.

I would like to thank **Dr. Amitava Patra**, Director, INST Mohali for his constant interactions with students and motivating students through his speeches. His message during one of his speeches, "To succeed in life, you don't need to be the best, but rather unique," has made a lasting impression on me.

I thank Central Analytical Facility, INST Mohali for making all the facilities available required to carry out the research. Sincere thanks to the administrative staff for timely help in fellowship and academic-related issues. Special thanks to the purchase section for being kind and helpful during the procurement of consumables and non-consumables. I would also like to thank the non-administrative staff for ensuring the smooth functioning of daily activities.

I would also extend my gratitude to **Dr. Sharvan Sehrawat**, IISER Mohali, for allowing me to use their lab facilities whenever needed.

The journey as a PhD research scholar is full of ups and downs, some days cheer you up with successful experiments, and some force you to think beyond your imagination and push your limits. It becomes easy when you have wonderful people around you as lab-mates. I sincerely thank **Dr. Naimat Kalim Bari**, the most experienced and helpful person in the lab. He introduced me to the lab, instruments, and all the ongoing research in the lab. I got a chance to work with him and I learned how passionate he is about his work which always encouraged me. I will always appreciate his kind and generous help whenever needed whether in making me understand the complex macromolecular system, fascinating nanoparticle world, experiments, or making presentations. **Ankush** was working on a different research problem of p53 protein. She had a jolly nature and I am glad we could connect later with our ongoing love-hate friendship. **Gaurav,** I always admired your scientific temperament and practical approach. Discussions with you, whether scientific or personal, have always given me direction. Our trip to Jim Corbett National Park will always remain memorable, as it has become a r source of laughter for lab get-togethers. **Dimple** joined the lab when I was in the third year of my Ph.D. journey. She has been a great support during the last three years which I cannot put in words. I always felt happy teaching you because you are a quick learner. If I have to ask for any help I have always looked upon you because I have confidence in you and I know that it will be done. **Preeti,** I always admire you for

your positive and independent approach. Your creative and joyful nature fills up the lab. Of course, you don't like but you have incredible management skills. Thank you for supporting and understanding me. **Silky**, there were many ups and downs but I acknowledge your help whenever needed and cherish happy times. **Harpreet**, I have always enjoyed our fights over petty issues. They were entertaining and kind of a stress buster. I must say far along you came out to be the most understanding person. Rose and Aarcha, you joined last to the lab and I am grateful to them for the love, kindness, support, and encouragement, I received. Rose is a person with a clear mind and Aarcha is confident and hard-working.

I am immensely grateful to SR lab members, **Surbhi, Nisha, Gayatri, Jagadish, Sai, Malay, Sayan, Pritam, Bhati, Veerpal, Tanuja, and Devansh** for maintaining a healthy atmosphere and always being supportive. I will always cherish the lab trips altogether.

I would like to extend my gratitude to **Dr. Charu Sharma,** CSIR-IMTECH, Chandigarh for allowing me to work in her lab. It was my first experience as a researcher. Her guidance and experience which I have gained in her lab have significantly helped me in my PhD research. I am very thankful for the knowledge and skills I gained while working in her lab.

I was fortunate enough to be supported by wonderful and loving friends throughout, whom I met from the research group and outside. I happened to meet **Surbhi, Kritika, Aabha, Ginni, Chirag, Neetu, Navi, Deepu, Vibhuti, Aman, Harmeet, Ashish, Kanu, Shashi, Subhash, Sarab, Gaurav**. They have constantly supported me throughout the journey. I cannot forget how **Kritika, and Aabha, you** had kept me motivated and positive all the time during the first year of my PhD when I was completely broken. Thank you for helping me survive that year. I still cherish our chats and gossip in LHC.

I am incredibly grateful to my friend **Surbhi**; your unwavering support has been indescribable. You selflessly helped and supported me through thick and thin. You have always been there in difficult times when I was most in need, and never let me quit. You were always there to encourage me and keep me going. Without your support, this journey would have been much more difficult. Thank you for making it smoother.

I cannot express my gratitude enough towards my **parents and my sister**, who have been the pillars of support during my journey and before. All their sacrifice and patience during this journey and before is beyond words which cannot be described. I never had to ask them to understand me or explain to them about not being available in many instances. Their love and support have been a constant source of strength for me I was lucky to be served with love all the time. I am grateful to my parents for always being proud of me. I have learned to be strong from my father and the art of patience from my mother which has helped me during my PhD. My **sister Jaspreet and my brother-in-law Jugraj** have always been super supportive by being there and being present in circumstances where I couldn't. I am grateful to my **Chachaji** and **Chachiji**; she has been a great support. You are more of a friend with whom I can share my deepest feelings. You understand me much better than anyone else. You loved me and supported me unconditionally during difficult moments. You are the reason I could come out easily through the toughest time of my life.
I would like to acknowledge all the support I got from **Amar**. You have your own intense ways but you have always given me a clear vision, encouraged me, and, supported me in many aspects.

Lastly, I thank everyone who has helped me in any aspect and made the journey manageable. This endeavor would not have been possible without the support of all.
Thank you for putting smiles on hard times.

Simerpreet Kaur

Abbreviations

1,2-PD	Propane-1,2-diol
2D	2-dimensional
3D	3-dimensional
4-AP	para-Aminophenol
4-NP	para-Nitrophenol
ARC	Antibiotic resistance cassette
Au-NPs	Gold nanoparticles
BMCs	Bacterial Microcompartments
BMV	Bacterial microcompartment vertex domain
BPER-II	Bacterial Protein Extraction Reagent
CD	Circular dichroism
CO_2	Carbon dioxide
CRISPR	Clustered regularly interspaced short palindromic repeats
DLS	Dynamic Light scattering
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DSF	Differential scanning fluorimetry
EUT MCP	Ethanolamine Bacterial Microcompartment
FRET	Fluorescence resonance energy transfer
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	Isopropyl β-d-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
LB	Luria broth
LCST	Lower critical solution temperature
MBTH	3-Methyl-2-benzothiazoline
$NABH_4$	Sodium borohydride
NCE	Non-carbon E salts media

PCR	Polymerase chain reaction
PDU MCP	1,2-propanediol utilization microcompartment
PMSF	Phenylmethane sulfonyl fluoride
RFU	Relative fluorescence unit
RMSD	Root mean square deviation
SDS-PAGE	Sodium dodecyl sulphate - Polyacrylamide gel electrophoresis
SO	Sypro orange
TEM	Transmission electron Microscopy
TEMED	N,N,N′,N′-Tetramethyl ethylenediamine
TMB	3,3′,5,5′-Tetramethylbenzidine
UV-Vis	Ultraviolet
Vit B12	Vitamin B12 (Ado/Cyano cobalamin)
WT	Wild type

Abstract

Understanding the fundamental principles behind macromolecular structure-function interactions requires an understanding of their complexity and dynamics. The complex organization of a system entails diverse components interconnected in different ways. Bacterial microcompartments (BMCs) are such types of sophisticated nano-machinery that involve unique protein assemblages of thousands of protein subunits forming mega Dalton (MDa)-sized icosahedral organelle in prokaryotes. They function to provide adaptive growth advantage in special environment conditions by segregating segments of metabolic pathways which include toxic intermediates. The compartment's structural design contains a catalytic core surrounded by selectively permeable shell proteins which belong to the conserved BMC domain family. The canonical domains present are either single–domain or tandemly fused dimer domains, oligomerizing as homohexamers or pseudohexamers with a comparable conformational topology with the α/β BMC domain fold. Over the evolution, the divergence in amino acid sequences and genetic occurrences in BMC have resulted in complicated protein fold variants. The domains are also associated with circular permutations emphasizing the structural variety of shell proteins. Thus, much diversity is embedded into them which is underexplored. My thesis strained the question of sequence and structural diversity in the shell protein domains. In this context, I have used biophysical approaches to get insights into this puzzle by exploring one of the most complex metabolosomes, 1,2 propanediol microcompartment (1,2 Pdu MCP). The outer shell of 1,2 PduMCP is a mosaic of eight different shell proteins and is encoded by a single operon. I have used two key shell proteins of the 1,2 Pdu MCP, PduA and PduB' to create synthetic shell protein variants by *in vitro* domain swapping. Two approaches were used, in one instance, synthetic dimer proteins were created by fusing two PduA domains to produce a PduA dimer that is circularly permuted and resembles PduB'. In this work, the potential role of domain dimerization (pseudohexamer) and therein circular permutation is addressed. Next, the PduB' dimer was broken down into constituent domains as single-domain proteins, and a circularly permuted variation of PduA was created to be compared to PduA. Apart from *in-vitro* studies, I have also substituted PduA protein chromosomally *in-vivo* with its circularly permutant variant. My research investigated the significant role and effect of structural topology and their exquisite sequence specificities. These

new protein variants with native sequences shed light on the evolutionary significance of various sequences, the structure of shell proteins, and how these factors relate to BMC's overall stability.

Synopsis

Evolution takes shape towards the compex organization of living forms. The complex organization demands many components to interact and function in a coordinated manner. This makes them resilient and offers good benefits towards changing environments. Recent advances have demonstrated that bacterial cells like their eukaryotic counterparts have a presence of complex organization of organelles known as Bacterial microcompartments (BMCs). These BMCs are sophisticated nano-machinery that involve a unique assemblage of thousands of protein subunits and form a mega Dalton (MDa)-sized polyhedral organelle. They are conditionally expressed in prokaryotes and appear in a 100-200nm size range. Many species of bacteria express these organelles which serve the purpose of compartmentalization for enhanced efficiency and segregation of metabolic processes which are otherwise detrimental to the cell due to toxic intermediates. The structural architecture of BMC contains a catalytically active enzymatic core surrounded by selectively permeable shell proteins. These shell proteins belong to the conserved family of bacterial microcompartment domain proteins (Pfam00936). The BMC domain forms the basis of the shell proteins. Their self-assembly leads to two different types of oligomerization, i.e., hexamer and psuedohexamers. Hexamer is formed from the self-assembly of six (single-domain) canonical BMC domains and psuedohexamer is formed from two fused BMC domains (dimer-domain) which self-assembles to form a dimer-trimer. Both types of oligomerization form structurally equivalent proteins with a flat hexagonal symmetry that tile together to form the facets of the BMC polyhedral structure. It is exciting and equally challenging to understand such a proteinaceous organelles due to its structural complexity. Since their discovery, they have attracted a lot of attention and have been a fascinating subject of study due to their nano-biomaterial and nano-biotechnological applications.

The present thesis attempts to decipher the structural complexity of shell proteins. Over the evolution, the variations at the core of the BMC domain at the level of divergence in amino acid sequences and modular rearrangements in the structural domains have resulted in complicated fold variants that offer distinct properties and functions. The domains are also associated with circular permutations emphasizing the structural variety of shell proteins. This thesis strained the question of different oligomerization and, sequence and structural diversity in the shell protein domains. In

this context, to get insights into the puzzle, I have explored one of the most complex metabolosomes, 1,2 propanediol microcompartment (1,2 Pdu MCP) from *Salmonella enterica* LT2. The bacteria produce 1,2 Pdu MCP which provides a growth advantage by utilizing 1,2-propanediol (1,2-PD) as an alternate carbon source in energy dearth conditions. The outer shell of 1,2 PduMCP is a mosaic of 8 different shell proteins which are encoded by a single operon. The operon has 23 genes; pocR, PduF, and PduABCDEGHJKLMNOPQSTUVWX, where 8 genes code for the shell proteins i.e. PduA, B/BB', J, K, M, N, T, U) and other codes for enzymes involved in degradation pathway and cofactor recycling. These genes for all the shell proteins and enzymes are clustered together in an operon in a sequence-specific manner which determines the biogenesis of BMC and makes it difficult to decipher the functional role of individual shell proteins. Among the 8 shell proteins, I explored PduA and PduB' in my thesis which occupy the first and the second position in an operon respectively, and make the major composition of the outer shell which oligomerize as hexamers (single-domain; PduA) or pseudo-hexamer (dimer-domain; PduB') from the conserved BMC domain.

In ***Chapter 1***, I begin by introducing the complexity and diversity of living forms that benefit the organism to become more efficient and resilient to changing environments. Subcellular compartmentalization is a feature of complex organization of eukaryotes which is characterized now in prokaryotes as well. I have described briefly the discovery and evolution of these compartments in bacteria which are widely called bacterial microcompartments (BMCs). These compartments are different from the organelles present in eukaryotes. Based on their function BMCs are broadly categorized into anabolic BMCs and catabolic BMCs. The primitive anabolic BMCs are involved in CO_2 fixation and are hence called Carboxysomes. Catabolic BMCs are also called metabolosomes and are involved in the metabolism of various organic compounds such as propanediol and ethanolamine. Metabolosomes provide a growth advantage to bacteria in certain energy-dearth conditions by encapsulating the toxic intermediates in the metabolic pathway. The semi-permeable protein shell encases the enzymatic core within the lumen and separates from the cytoplasm. A description of carboxysomes and metabolosomes is provided. Further, I have described the structural architecture and composition of the BMC shell. The intrinsic self-assembly property of the shell proteins makes them unique supramolecular protein organelles. The shell proteins belong to the conserved bacterial microcompartment domain family which forms the

facets of BMC polyhedral structure. A detailed description of the different types of shell proteins forming the basic building block of its architecture and different oligomerization states of shell proteins in BMCs is given. Here I explain the different oligomerization of BMC domains as cyclic hexamers and psuedohexamers. Further describing how BMCs evolved to more complex and diverse structures I moved on to the research theme of my thesis.

A detailed introduction about 1,2 Pdu MCP is given focusing on the structural complexity. The 1,2 Pdu MCP shell has 8 different proteins. I have given information about them focusing on single-domain, dimer-domains, and circularly permuted single and dimer-domain proteins. Followed by research questions about the existence of oligomerization in shell proteins as hexamers and psuedohexamers, sequence and structure diversity, circular permutations in shell proteins, the potential role of domain dimerization in providing stability and flexibility, and its implications in the global function of microcompartment.

Chapter 2 provides information on various methods, techniques, and protocols used in this thesis. I have used a barrage of molecular biology and biophysical techniques. The new methodology and protocols used are mentioned in detail in this chapter. All the materials used, newly generated gene constructs *in vitro,* and newly developed strains *in vivo* are given in tabulated form.

Chapter 3 addresses the first question of shell protein complexity. I have compared the existence of two domain types and in this chapter potential role of domain dimerization is addressed. It is also related to the question of sequence and structure variations. BMC contains single-domain proteins or dimer-domain proteins. Specifically, 1,2 Pdu MCP has two types of native dimer-domain proteins, PduB' and PduT. PduB' has two fused domains and the domains are circularly permuted variations of canonical domain fold. PduT is a result of a direct duplication event where two domains are fused in a tandem fashion. Apart from being related through circular permutations, they have a variance in sequence. To get insights into this puzzle of variance in sequence and structure, the native sequence of PduA single-domain protein is explored, and a dimer of PduA and circularly permuted dimer which is structurally similar to PduB' was generated. I have exploited molecular biology techniques to clone designed constructs and overexpress *in vitro* to understand the structure-function relationships in the BMC stability. To understand the stability and differences in the newly designed synthetic proteins which carry a sequence from

PduA protein and structure similar to native dimer-domain proteins. I have used biophysical techniques like circular dichroism. fluorescence spectroscopy to probe the folding of the proteins. Dynamic light scattering studies have shown the proteins retain the self-assembly property. However, TEM micrographs show that all different protein variants generated self-assembles to provide different architectures. Further temperature-dependent stability studies provide insights into the important role of sequence in governing biophysical properties. We found native dimer proteins' secondary structure elements uncoil at above 80^0C and regain the structure post-melt from circular dichroism spectroscopy. Thus pointing toward the flexible nature of native proteins. In contrast to synthetic dimer proteins which don't show uncoiling of secondary structure elements even at 95^0C. This characteristic is similar to the parent protein PduA. Irrespective of the topology attained the properties belong to the original core of the amino acid sequence. Hence, protein sequence governs the biophysical properties of BMC domain proteins and consecutively their functions as well.

In ***Chapter 4*** another interesting question about the existence of single-domain shell proteins in BMC is addressed. In chapter 3 what was found about the relatedness of shell proteins by sequence and structure is interesting. Therefore, in this chapter, I intend to get insights into the probable role of single-domain proteins in the global stability of BMC. The question here is what if the sequence from the dimer domain protein is explored to generate a single-domain protein? The study of their properties will be interesting when compared to native single-domain proteins. So, in this line, single-domain variants were generated. PduB' protein was dissected into its individual domains from the linker region joining the two domains to get domain1 and domain2 separately. The existence and their probable role as individual domains are then compared. I have also explored the sequence from PduB' to generate a circularly permuted single-domain protein having a topology similar to PduA. For this purpose, I performed structural alignment of PduA and PduB'. The alignment shows PduA aligns to the second domain of PduB' with an RMSD score of 0.85Å than the first domain having an RMSD of 1.49 Å. So, I chose the second domain and circularly permuted it to make it like PduA. All three single-domain protein variants were generated and purified *in vitro* using molecular biology techniques. Protein variants were folded in their native conformations after purification as assayed by circular dichroism and fluorescence experiments. Further studies by DLS and TEM, evidently proved that they retained the self-assembling property

and formed higher-order supramolecular structures except PduB' domain1. Temperature-dependent results show that the single-domain protein variants are thermostable similar to PduA. CD thermal melt shows that the secondary structure does not unfold to 95°C thermal ramp. The individual domains don't show a flexible nature as observed in native PduB', thus this study suggests that this dynamic behavior is due to the presence of both domains together which governs the overall flexible nature and thus, provides resilience at a wide range of external stimuli.

In ***Chapter 5*** of this thesis, I have developed *in vivo* mutant strains where the first two positions of the Pdu operon, PduA and PduB' have been explored. The new synthetic protein variants generated *in vitro* in chapters 3 and 4 with native sequences have provided insights into the evolutionary role of diverse sequences and structure of shell proteins and their implications in the global stability of BMC. In this chapter, I incorporated one of the circularly permuted protein variants of PduB' domain2 to be substituted chromosomally in place of its native counterparts i.e., PduA. PduA occupies the first and most important position in an operon and serves as a scaffold for interaction with other proteins. Furthermore, PduA gene position is switched by the second gene PduBB'. So a new mutant strain is generated which has PduBB' gene at the first position and PduA occupies the second position. For developing *in vivo* strains I explored and standardized a new powerful tool, CRISPR Cas9 genome editing. The shell proteins tend to self-assemble as a sheet in 2-D due to interaction among domains and shape complementarity which forms the facets of polyhedral structure. Due to the structural similarity of shell proteins, deciphering the exact role of individual shell proteins is challenging. To determine whether the shell protein replacements have a substantial effect on the growth of the bacteria, the growth curve of wild type and modified strains in minimal media was evaluated in the presence of an inducer, 0.6% 1,2 propanediol (1,2 PD) supplemented with 150 nM CN-B_{12} cofactor as reported earlier for growth assays. The results showed a longer lag phase and delayed growth of the modified strains than wild wild-type variant. Further with many attempts, we could not successfully purify the microcompartment from mutant strains. This suggests that PduA protein substitutions have a vital impact on the structure and function of MCP where it abrogates the proper MCP formation and hence leaky MCP could not mitigate the toxicity from volatile intermediate, propionaldehyde.

Chapter 6 In this chapter, I have explored PduA shell protein for material science application for the development of hybrid nanocatalysts. The most stable and well-studied shell protein PduA and its mutant PduA[K26A] exhibit two different architectures. Their reported crystal structures reveal the formation of stable well-ordered 2D protein sheets by PduA (PDB ID 3NGK) while its mutant PduA[K26A] (PDB ID 4QIF) where lysine was mutated to alanine which is crucial for interactions in forming 2D sheets alters its morphology to aberrant 3D supramolecular architecture. I explored 2D and 3D architect scaffolds for the fabrication of gold nanoparticles. These naturally self-assembling proteins in 2D and 3D architectures reduced gold salt to gold nanoparticles and result in the different spatial organization of gold nanoparticles. It is also seen from TEM images and histogram plots that the varying architectures resulted in different sizes of nanoparticles. The average particle size distribution of PduA[K26A]_NPs is 4.0 ± 0.07 nm with a narrow distribution from 2.5 nm to 5.5 nm whereas PduA_NPs range from 5.5 to 12.5 nm with an average particle size distribution of 8.4 ± 0.1 nm. Further protein nanoparticle conjugates when explore for catalytic efficiencies it is observed a two-fold increase in catalytic activity for the reduction of 4-nitrophenol. They have shown to have remarkable peroxidase mimicking catalytic activities. The six-fold higher affinity toward TMB (3,3′,5,5′-tetramethylbenzidine) substrate is observed as a measure of higher peroxidase activity (nanozymatic) in the case of PduA[K26A] 3D scaffold. This approach demonstrates the understanding of scaffold architecture on nanoparticle synthesis and properties, thus paving the way toward the design of new nanomaterials.

Chapter 7 summarizes the key findings of the research work presented in this thesis. This work brings some insights into the puzzling question of BMC shell protein complexity. Over the course of evolution, the conserved BMC domain, a single building block from the Pfm00396 BMC domain family has resulted in sequence and structure variations. This diversity is not seen in another pentamer family (Pfam03319). Thus these variations have some important implications. This thesis sheds light on the implications of single and dimer domains in the global stability of BMCs. Single-domain proteins forming hexamers are more rigid and compact in structure in comparison to dimer-domain proteins which form pseudohexamer oligomerization. Native dimer proteins PduB' and PduT show the unfolding and uncoiling of secondary structure elements at high temperatures which they regain post-melt. Hence, they are dynamic structures that provide the BMC with flexibility, and proteins like PduA provides rigidity to the structure. Further, when

accessed for the role of sequence and structure, though, dimers from the PduA sequence resisted unfolding like PduA, irrespective of the topology attained the properties belong to the original core of the amino acid sequence. Thus, the sequences in native dimer proteins provide the desirable characteristics of the proteins for their function. The dissection of individual domains from PduB' has deciphered their probable role. The single domains of PduB' protein work in a synergistic manner when present together. Dissected domains were found to be more compact whereas together they exhibit dynamic structure. It is also evident by *in vivo* experiments. Studies by substituting the PduA shell protein with a structurally similar protein but with a different sequence abrogated the proper MCP formation. Overall chapter 3-5 concludes that oligomerization and their exquisite sequence specificities play an important role that balances and provides both rigidity and flexibility to MCP to function better in a coordinated manner. Also, synthetic protein variants have shown remarkably different higher-order structures. PduA single domains when fused as PduA dimer, it forms tubes in contrast to PduA which forms 2D sheets. On circularly permuting the second domain of PduB' its ability to form long thin sheets reduces and it instead forms small sheets. Thus these protein variants can be approached for biotechnology and material science applications. Chapter 6 delivers an approach towards application of shell proteins in material science. It provides an insight into precise control of the hierarchical organization of shell protein-nanoparticle conjugates in 2D and 3D to develop nanoscale materials with tunable properties and catalysis.

Appendix

In the appendix part of this thesis, I have investigated the potential of BMC to be developed as a thermoresponsive protein-polymer nanoreactor. Because of their proteinaceous shell, BMC can be targeted via extra functionalization using proteins, antibodies, polymers, etc. They can be used as a synthetic metabolic hub for heterologous cargo loading for various applications. In this part of the work, I aimed at designing a BMC-polymer nanohybrid by extra-functionalization using thermoresponsive polymer (Poly N-isopropyl acrylamide). Thermoresponsive polymers show a property of Lower critical solution temperature (LCST), above this critical temperature they tend to collapse and aggregates by excluding water molecules where we aimed to block the pores present in the center of the hexamer building blocks which are conduits for substrate and cofactor

diffusion in and out of BMC. We hypothesized combining of using this property to tune the activity of BMC which will be probed by using the CDE (diol dehydratase) enzyme present within the BMC lumen. Polymer was characterized by gel permeation chromatography, dynamic light scattering, and absorption spectroscopy. This provided us with the LCST temperature of the polymer. Further, we used Schiff base reaction to conjugate the aldehyde functionalized polymer to the BMC exposed basic lysine amino acid residues. We characterized the conjugate by DLS and SDS-PAGE. On assaying the temperature dependent activity, above and below LCST, we found that the change in BMC activity corresponding to the concentration of the polymer used and little but significant effect towards temperature. We used different concentrations of polymers and explored different conditions to successfully get the thermoresponsive nanoconjugate.

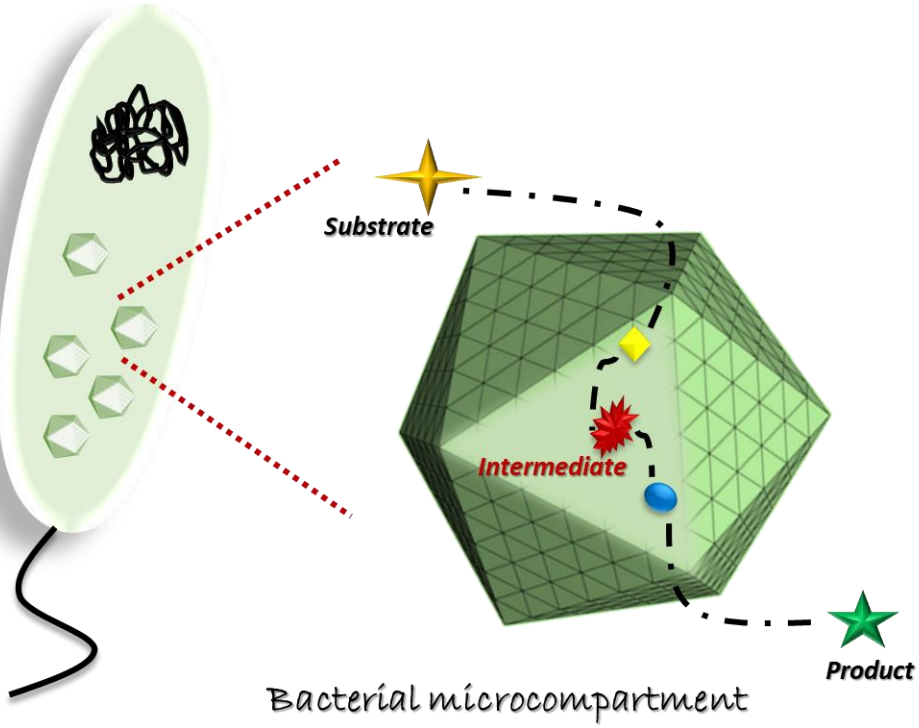
Table of Contents

List of Figures

List of Tables

CHAPTER – 1

Introduction



Bacterial microcompartment

1.1 Bacterial microcompartments: nano-machinery in prokaryotes

Prokaryotes have continuously developed ingenious strategies to thrive in challenging and unforgiving environments, adapting to their advantage. Over several decades, a multitude of discoveries have unveiled the existence of geometric-specific organelles within bacterial cells. The existence of compartmentalization as lipid-bound organelles is well established in eukaryotes but breaking the old notion, 70 years ago, prokaryotes were discovered with the occurrence of unique geometric-specific protein compartments. They are widely called Bacterial microcompartments (BMCs). Bacterial microcompartments are tiny polyhedral-shaped protein machinery that houses specific biochemical reactions by encapsulating enzymes within a protein shell, providing the spatial organization and segregation of metabolic processes within a cell. This compartmentalization serves several advantages to the bacteria. It is critical for efficient enzymatic reactions by accumulating and concentrating substrates and enzymes and preventing unnecessary cross-talk with other undesired molecules or reactions. They also protect the bacterial cells from the detrimental effects of toxic intermediates produced in certain metabolic pathways by enclosing and segregating that particular portion of a reaction. The selectively permeable protein shell is a key feature of BMC physiology which controls the diffusion of specific substrates, cofactors, and products in and out of the compartment. It prevents the exit of essential substrate molecules in sequestered metabolic reactions, and seized toxic metabolic intermediates are metabolized and converted to non-toxic compounds that are subsequently allowed to exit the BMC. They have evolved to different environmental conditions and metabolic requirements in different organisms. The most interesting characteristic is that despite being involved in different reaction mechanisms in different bacteria, it follows the same principles and building blocks to assemble similar outer shell architecture to house different enzymes and reactions.

It is seen that there is a lot of diversity among BMC types which captivated my attention. Despite being constructed from similar building blocks; BMCs have shown a complex organization of outer shell by assembly of thousands of protein subunits to encapsulate enzymes within. The genes associated with microcompartment proteins (BMC genes) are typically encoded within a single operon, although their length may vary, ranging from 10 to 12 kilobases (kb), or they might be

scattered across independent loci within the genome. It is quite demanding to understand how a complex and organized system is assembled and functions.

1.2 Microcompartment Types

The micrcocompartments are typically classified into different categories based on the types of biochemical reactions they participate in, specifically whether they are anabolic or catabolic in nature[1–3]. Anabolic BMCs have a significant role in global CO_2 fixation by encapsulated enzymes, RuBisCo, and carbonic anhydrase by carbon concentrating mechanism within a lumen[4,5]. Hence, named Carboxysomes. Catabolic BMCs collectively called metabolosomes, harbor substrate-specific signature enzymes within the lumen, for example, propanediol dehydratase or ethanolamine ammonia-lyase which are involved in the degradation of their corresponding substrates, propanediol or ethanolamine[6,7] (Figure 1.1). They are named based on the type of substrate metabolism they are involved in, propanediol utilization (Pdu) or ethanolamine utilization (Eut) micrcompartments (MCPs). Their degradation leads to a toxic aldehyde intermediate which is further metabolized to alcohol and carbonic acids[8,9]. The discovery of these microcompartments occurred after the identification of shell protein genes that exhibited homology with carboxysomes[10]. The shell protein genes play a crucial role as markers in the quest to identify potential microcompartment loci in prokaryotic organisms. The advent of genome sequencing data and advancements in bioinformatics tools have significantly contributed to the discovery of new classes of microcompartments, primarily through the identification of these distinctive shell proteins. It is reported to have found BMC loci in 45 bacterial phyla and identified 68 BMC types/subtypes[11]. The knowledge has expanded in the case of the existing type of metabolosomes from propanediol and ethanolamine utilization MCP to other identified that involve the metabolism of ethanol, choline, aromatic amino acids, sugar phosphate, and aminophenol[11–18].

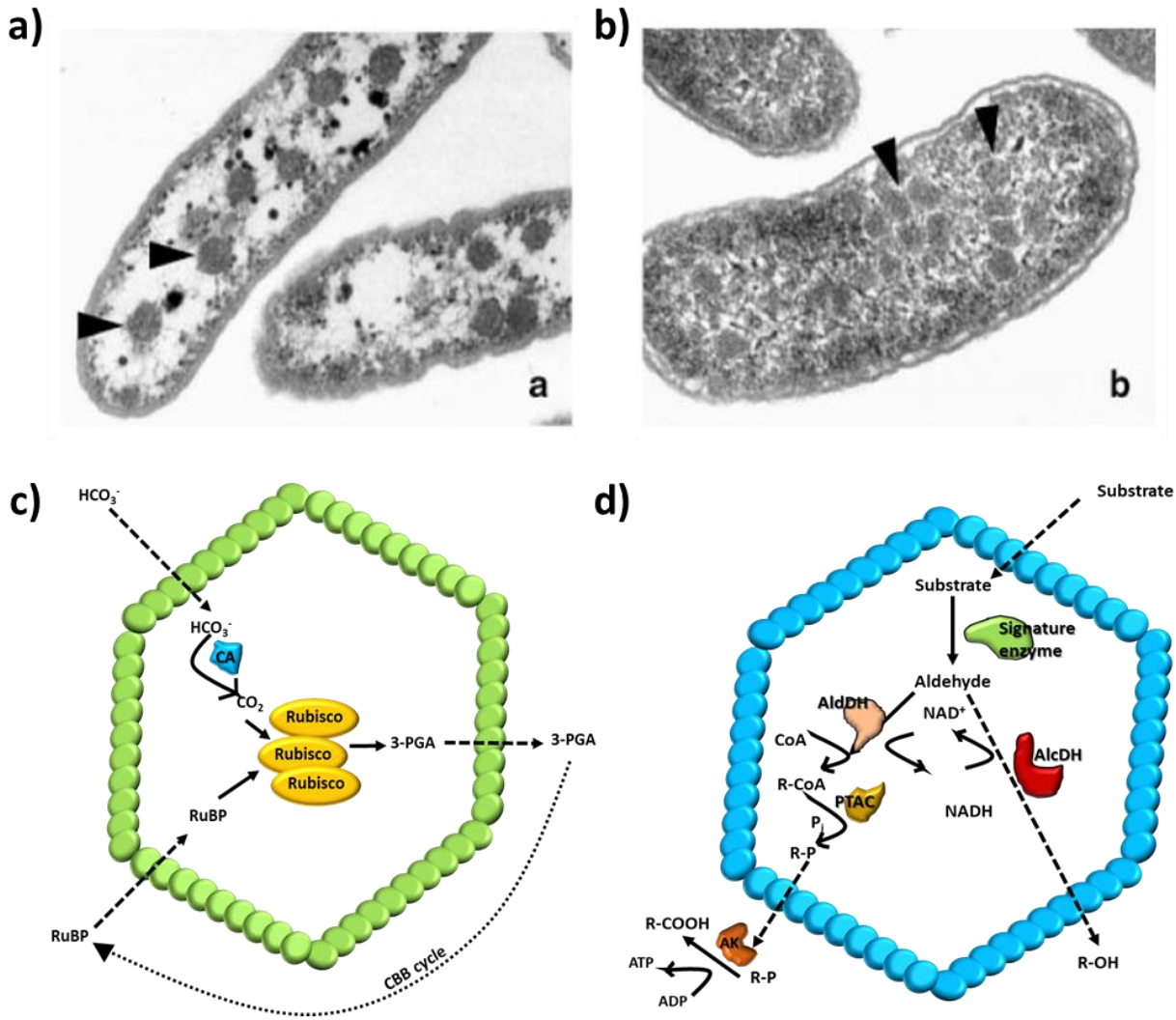


Figure 1.1 TEM images of bacterial microcompartments. a) Thin-section micrographs of carboxysomes in *Halothiobacillus neopolitanus*. b) Thin section micrographs of Pdu metabolosomes in *Salmonella enterica* (adapted from *Cheng et al.*[19]). Schematics representing c) carboxysomes d) metabolosomes (adapted from *Yeates et al.*[20]).

1.2.1 Carboxysomes

Carboxysomes were first observed in cynobacterium *Phormidium uncinatum* in 1956, as intracellular structures[21]. They were initially thought to be phages due to their structural resemblance to viral capsids. Later, in 1969, similar structures were found in many blue-green algae and were seen as polyhedral crystalline bodies with highly ordered arrays through electron micrographs[22]. In 1973, Shively et. al, isolated these polyhedral structures from the chemoautotrophic bacterium *Halothiobacillus neapolitanus* and discovered to contain an enzyme ribulose bisphosphate carboxylase (RuBisCO)[4,23]. They are then known to participate in carbon dioxide fixation as part of the Calvin-Bensen-Bassham cycle[24]. RuBiscCo has an affinity both for oxygen and carbon dioxide. It has, however, less affinity towards CO_2. The carboxysomes therefore combine the carbonic anhydrase enzyme with RuBisCo within a shell which converts then bicarbonate into CO2. It concentrates them together in a shell and prevents the CO_2 escape to increase the catalytic efficiency of RuBisCo enzyme[25].

There are two types of carboxysomes depending on the type of RuBisCo and carbonic anhydrase[26]. The alpha-type carboxysomes are associated with Form IA RuBisCo and are found in chemoautotrophs and some marine cyanobacteria. The genetic information for alpha carboxysomes is organized into discrete operons and termed as cso (car-boxysome) operon[27]. The beta carboxysomes are associated with Form IB RuBisCo, found in freshwater cyanobacteria. They are encoded by ccm (carbon concentrating mechanism) operon and are distributed throughout the genome[28,29]. They are about 100-200 nm in diameter and are found to have more regular icosahedron structures.

1.2.2 Metabolosomes

Metabolosomes are different from carboxysomes in structure and function where they are involved in catabolic reactions[19]. Homologs of carboxysomes shell proteins were found and many types of metabolosomes have been discovered and the list is going on[30]. These microcompartments play a fundamental role in the metabolism of organic substrates and are named based on the specific type of substrate they handle. For instance, two of the most extensively studied microcompartments are

the propanediol utilization microcompartment and the ethanolamine utilization microcompartment, which metabolize 1,2-propanediol and ethanolamine as their respective substrates. One common trait among these microcompartments is their ability to shield the cell from DNA damage caused by volatile and toxic intermediates that arise within the corresponding metabolic pathways. Consequently, they not only serve as a means to concentrate enzymes and substrates but also act as protective shields for the cell. In terms of structure, metabolosomes exhibit a more irregular configuration compared to carboxysomes and are typically found in closer proximity to beta carboxysomes. Their size typically falls within the range of 100-150 nanometers[31].

1.3 BMC shell protein domain and oligomerization

The outer covering of the BMC structure is constructed from three types of basic building blocks BMC-H, BMC-T, and BMC-P [32–36]. Hexamer shell proteins are termed BMC-H, pseudohexamer shell proteins are termed BMC-T, and pentamer proteins are termed BMC-P (Figure 1.2). The outer shell of BMCs is constructed by these same architectural building blocks and this is a common shared structural feature among all BMCs. The shell protein building blocks belong to a conserved protein domain fold from the bacterial microcompartment (BMC) domain family (Pfam00936) and bacterial microcompartment vertex (BMV) domain family (Pfam03319)[1,37] which is a reason that the shell proteins are highly conserved despite the functional diversity in BMCs. The presence of the BMC domain confirms that the protein assembles into a hexameric structure, while the presence of the BMV domain forms a pentameric assembly. The variation in a BMC domain which can occur either as a single instance or as a tandem repeat within the protein sequence, results in the formation of either a hexameric or a pseudo-hexameric protein assembly. The hexameric and pseudo hexameric proteins exhibit a comparable overall conformation, resembling a flat, hexagonal-shaped protein structure. This structural similarity enables them to spontaneously assemble randomly into a two-dimensional molecular layer, forming the facets of the BMC shell. Meanwhile, the pentameric proteins take up positions at the vertices, contributing to the closure of the icosahedral structure. BMC shell proteins are found to be an adaptation from PII signaling protein and OB-fold domain-containing proteins from structural and sequence analysis[38]. It is suggested that these ancient proteins might have coevolved to diverse BMC shell

proteins as a result of parallel co-evolution. The widespread occurrence of BMC domain proteins across the bacterial kingdom suggests BMC genes have spread by horizontal gene transfer[39].

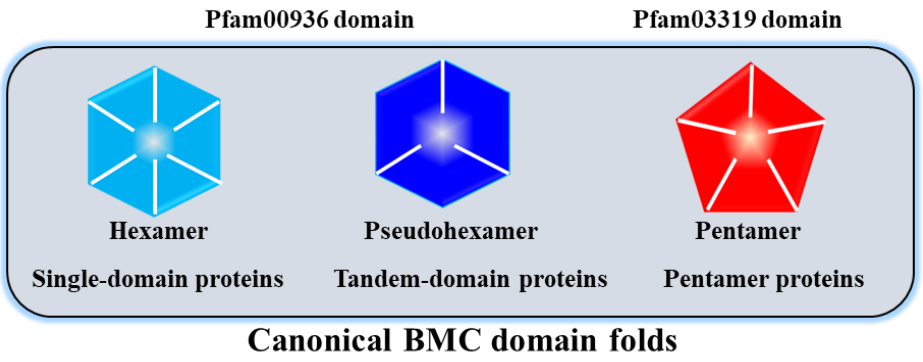


Figure 1.2 Schematics representing of canonical BMC domain folds forming a hexagonal and pentagonal symmetry. Hexagonal symmetry is shared by a hexamer and a pseudohexamer shell proteins; pentamer proteins adopts a pentagonal symmetry. (adapted from *Yeates et al.*[32])

1.3.1 BMC-H

BMC-H forms the major structural component of microcompartment architecture. Single domain (Pfam00936) make up the BMC-H proteins where six monomer units of a single domain oligomerize and give rise to a hexagonal disc shape assembly structure forming a pore at the central axis responsible for molecular transport across the shell[1]. Generally, one side of the hexamer assembly appears flat and polar whereas the other side bears a bowl-shaped depression having more hydrophobicity which is suggested to face the luminal side of BMC[40]. Another variation or subclass of hexamer proteins exists as circularly permuted hexamers (BMC-H^{P})[11,41]. The available crystal structures of hexamer shell proteins enriched us with information on their self-assembly and interacting residues responsible for it[1,42–44]. The shape complementarity of hexamers allows the side-by-side association and formation of 2D molecular layer[45,46].

1.3.1 BMC-T

BMC-T forms another type of shell protein which features the presence of tandem repeat of BMC domain and oligomerize to give a pseudohexamer (dimer-trimer). BMC-T shares structural similarities with BMC-H as both form hexagonal-shaped discs. However, these proteins break the strict hexagonal symmetry exhibited by hexamers. They are also reported to have variations with circular permutations and duplications[47–50]. Like hexamer proteins, they also assemble from side-by-side and form a molecular layer for facets of an icosahedral shell. Other proposed models by crystal and solution studies suggest that double-layer shells are present in BMC where hexamer and pseudohexamer are observed to be present as stacks forming double-disc structures [46,51–53].

1.3.1 BMC-P

BMC-P forms a pentamer protein by assembly of five monomer units of another BMV domain (Pfam03319). BMC-P shell protein is crucial in attaining the closed polyhedral structure of BMC and is believed to be present at the vertices. Crystal structures of CsoS4A and CcmL from the α- and β-types of carboxysomes[36] showed their pentameric assemblies well-suited to be vertices in an icosahedral shell, which earlier remained unidentified due to low abundance. Other homolog proteins of the pentamer family are EutN, PduN, and GrpN present in different BMC involved in respective metabolic functions[36,54]. The genetic experiments confirmed their structural role where pentamer proteins were deleted and compartments were reported to lack the ability to form closed icosahedral structures [31,51,55,56].

1.4 Protein fold variations

The canonical BMC domain adopts an α/β –fold with a four-stranded antiparallel sheet flanked by small helices to about 90 amino acids in length [20]. The crystal structure of various BMC proteins has revealed remarkable variations (gene fusions and rearrangements) in a canonical type domain resulting in conformational and topological variations. Some BMC proteins are related through

circular permutations where secondary structure elements occur in different order retaining a similar tertiary structure. This rearrangement results in different spatial organization of N and C terminals which are suggestive of important implications in the structure and function of BMCs. For example, PduU is a circularly permuted variant of PduA canonical fold[41]. The rearrangement and occurrence of a new extended N terminal allowed it to form a six-stranded beta barrel which leads to blocking a central pore. Another example of single domain variation is EutS which is different in itself as it forms a bent structure at approximately 40^0 in hexagonal disc rather flat disc-shaped protein[57]. In addition to single-domain protein fold variations, the structural diversity is seen in tandem BMC domains as well. For example, a carboxysomes shell protein CsoS1D[51], EutL[47,57], EtuB[50]. They are reported to have duplications of permuted domain folds resembling PduU. PduT is a result of a tandem fusion of two canonical domains[42]. PduK protein contains an extra domain and helix turn helix structure indicative of a potential role in nucleic acid binding. Thus all these variations in protein folds and divergence at the level of amino acid sequences have made the understanding of BMC difficult.

1.5 Research paradigm

1,2-propanediol (Pdu) utilization microcompartment is the most studied metabolosome. After the discovery of carboxysomes, genomic analysis and identification of a PduA homolog of carboxysomes shell protein led to their discovery. Later polyhedral structures similar to carboxysomes were observed in *Salmonella* when grown in the presence of 1,2 propanediol[39,58–60] . Genomic analysis of *pdu* gene cluster is shown to be present in many soil-dwelling and enteric bacteria[19,62]. Bacteria that produce the 1,2 propanediol utilization microcompartment have the remarkable ability to metabolize 1,2 propanediol as their sole carbon source. This compound is a byproduct stemming from the anaerobic degradation of rhamnose or fucose in the human gut. Interestingly, the presence of 1,2 propanediol, generated by the activity of human intestinal microbiota, triggers the activation of the *pdu* operon. This operon contains genes responsible for both shell proteins and enzymes.

Within the microcompartment, these shell proteins and enzymes come together through self-assembly, creating a protective microenvironment. In this setup, the shell proteins act as an envelope, while the enzymes are enclosed within the compartment. This organized micro compartmentalization is critical because the metabolism of 1,2 propanediol involves the production of a toxic intermediate called propionaldehyde. To mitigate the harmful effects of propionaldehyde, the microcompartment facilitates its conversion into propionyl-CoA and 1-propanol[31]. It's worth noting that propionaldehyde has been demonstrated to induce growth arrest and DNA damage in the bacterial cell. In addition to its metabolic role, the Pdu microcompartment is also associated with pathogenesis and provides a growth advantage to the bacteria by enabling the utilization of 1,2 propanediol as a carbon source. This adaptability contributes to the bacteria's success and pathogenic potential.

The *pdu* operon has 23 genes; pocR, PduF, and PduABCDEGHJKLMNOPQSTUVWX, for the shell proteins, enzymes involved in the degradation pathway, and recycling cofactors. The outer shell of Pdu MCP is a mosaic of eight different shell proteins i.e. PduA, B/BB', J, K, M, N, T, U. The presence of many types of shell proteins indicates a complex organization and each protein certainly has a significant role to play (Figure 1.3). The major shell proteins PduA, PduBB' are reported to form extended sheets and form facets of the Pdu MCP structure. PduN protein belongs to the pentamer family and is present at vertices. The major enzyme present in PduMCP is the diol dehydratase (PduCDE) which catalyses the conversion of 1,2- PD to propionaldehyde. It is subsequently converted to propionyl coenzyme A or 1-propanol by aldehyde dehydrogenase(PduP) and alcohol dehydrogenase (PduQ). Propionyl coenzyme A is then converted to propionate by propionate kinase (PduW) to generate ATP. There are other enzymes present that are involved in the reactivation of diol dehydratase and vitamin B_{12} recycling such as cobalamin reductase (PduS), adenosyltransferase (PduO), diol dehydratase reactivase (PduGH) and L-threonine kinase (PduX).

1.5.1 PduA Homohexamer (Single BMC domain)

PduA protein is present at the first position in an operon. It has a central role among all shell proteins as it forms the maximum interactions with other shell proteins such as PduBB', PduJ, PduN, PduU and is the major shell protein component of Pdu MCP[63]. The gene sequence is a single BMC domain and encodes a small 94-amino-acid protein. Apart from being a scaffold for protein-protein interactions, its first position in an operon is also crucial. On swapping PduJ with PduA in an operon, PduJ is shown to rescue its function[64]. In another study reversing the operon and placing PduA at the end position in an attempt to develop a minimal compartment has failed to from regular BMCs[55]. In deletion studies, PduA single mutant is shown to form enlarged MCPs with an appearance similar to that of the wild type[61,65]. The crystal structure shows that the PduA hexamers are packed closely and form a continuous 2-dimensional molecular layer. PduA hexamer forms a selectively permeable pore at the central axis which is suggestive of 1,2-PD transport and restricts the efflux of propionaldehyde[34,66–68]. Structural and genetic studies showed serine S40 at the pore GSG motif is a major route of 1,2-PD transport and is controlled by the size and electrostatic properties. It is also shown by molecular dynamic simulations where diffusion of 1,2-PD is allowed by 3folds over propionaldehyde. PduA is demonstrated to be able to self-assemble into protein nanotubes when overexpressed in bacterial cells and invitro on purifications at high concentration[69,70].

1.5.2 PduBB'- Trimer (Pseudohexamer-Tandem BMC domains)

PduBB' is a pseudohexamer protein that is expressed from two overlapping genes in the same reading frame. It is present at the second position in an operon. The gene sequence is a duplication of the BMC domain in a tandem fashion. On expressing from alternative start sites it expresses a longer PduBB' and a short length PduB' protein[71]. PduB' and PduBB' are identical in amino acid sequence except that PduBB' is 37 amino acids longer in N-terminal extension than PduB'. The extra 37 amino acids are disorder regions which have established their role in binding to the enzymatic core[72–74]. Deletion studies showed that PduBB' deletion mutants are unable to form BMC s and hence is a critical protein[65]. It also oligomerizes into an overall similar hexagonal symmetry like PduA protein from a trimer of two fused domains. Similarly, PduBB' forms the extended 2-dimensional molecular sheets and participates in forming the facets of BMC. It is observed to form nanotubes when overexpressed at high concentrations *in vitro*[69]. Homolog

proteins similar in structure to PduBB' have been crystallized in two forms, open and closed conformation which is suggestive of gated pores for large molecules[51,57,75,76].

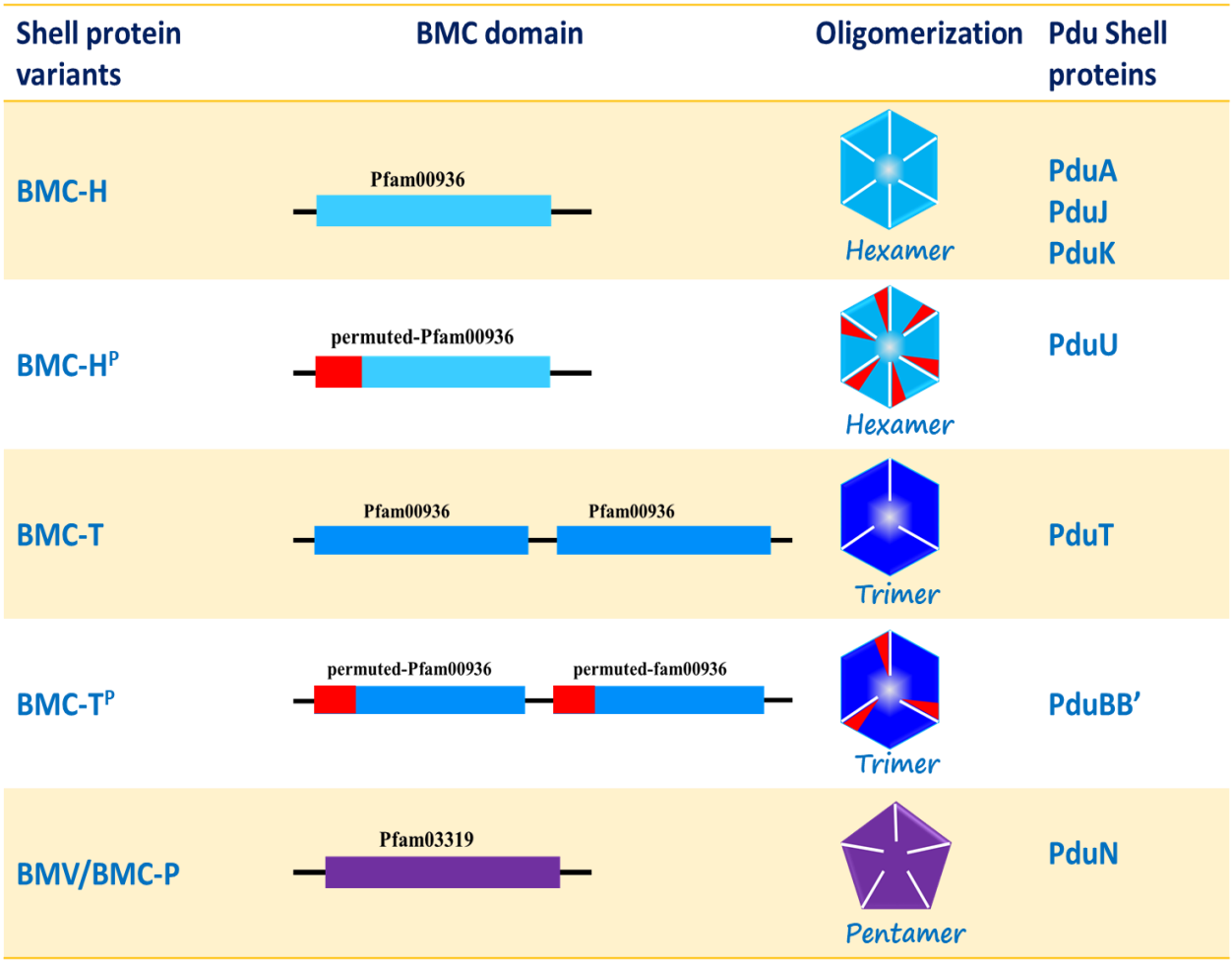


Figure 1.3 Protein fold variations making up the Pdu MCP shell. Eight different shell proteins with diverse domain folds adopting different architectures are shown. (adapted from *Yeates et al.*[72,77])

1.6 Research focus

Living forms have evolved towards complexity and diversity, although all organisms are derived from the same inimitable molecules that form the basis of life. Simple unicellular organisms comprise a single cell only which in itself is a complete organism and can perform all activities needed for life. More complex multicellular organisms comprise many different cells that acquire specialized roles, work independently, and all function in an integrated manner. The division of labor enables the cells to become more efficient. The diversity of living forms is the most intriguing phenomenon.

BMCs also exhibit a lot of structural and functional diversity. The variation in the overall architecture is witnessed between carboxysomes and metabolosomes. The BMCs are built from two types of shell protein modules hexagonal and pentagonal occupying the facets and vertices, respectively. This diversity is seen only in hexagonal building blocks and not the pentagonal ones. Thus they offer unique advantages to the microcompartment structure and function. Though the shell proteins encasing the lumen enzymes are phylogenetically related and belong to the conserved family of bacterial microcompartment domain proteins. It appears that the Pdu metabolosomes among other metabolosomes and carboxysomes are the most diverse and enriched with many types of shell proteins[78] (Figure 1.4). But one can wonder why a compartment needs such many different types of proteins making up the shell. What advantages does it offer with respect to structure and function? The current understanding of shell proteins is that during the course of evolution, the recombination and rearrangement events at the level of functional subunits have given rise to diverse BMC shells. The modularity in the shell proteins allows them to serve as bioplatforms for the development of nanoreactors, nanocarriers, or as hybrid scaffolds for harboring distinct catalytic reactions. Such applications put forward the potential of BMC to be explored for nanobiotechnology and material science applications. It is also easy to modulate BMCs because of their all-protein nature. To have such wide applications it is important to understand the fundamental building of BMCs.

The increasing complexity in the structure of microcompartments from carboxysomes to metabolosomes has enthralled my interest towards it. Further, the diversity in the hexagonal module in PduMCP is fascinating. Pdu metabolosomes comprise eight different types of shell proteins i.e., PduA, B/BB', J, K, M, N, T, U. The proteins forming the facets of the structure

comprise two types of self-assembled proteins oligomerized as hexamers and pseudohexamers. Hexamers self-assemble from single-domain proteins and pseudohexamer proteins self-assemble from tandemly fused dimer domains. The question I strained in the first research chapter of my thesis is why the microcompartment structure requires two types of oligomerized proteins. It could have been only hexamers from single domain proteins or only psuedohexamers from dimer domain proteins. What structural role do they serve in microcompartment stability? So, in chapter 3 of my thesis, I have addressed the question of the potential role of dimer domain proteins oligomerizing

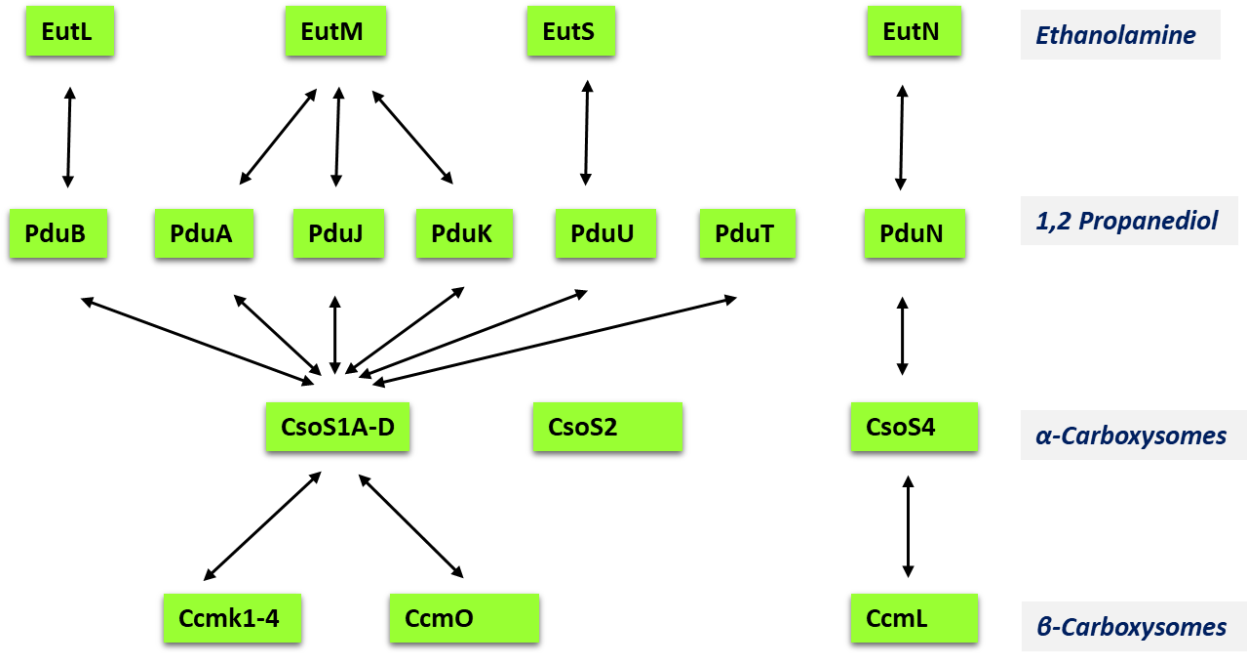


Figure 1.4 Similarity and increased diversity among shell proteins of carboxysomes and metabolosomes (adapted from *Frank et al.*[78]).

As pseudohexamers. Further, the implications of circular permutations and sequence on the structural stability of microcompartment is also addressed. For this purpose, I have used two native pseudohexamer proteins, PduB' which is a circularly permuted dimer domain protein, and PduT which is a direct tandem fusion of two domains. Another question is the role of hexamers obtained

from the self-assembly of single-domain proteins in structural stability. For this in Chapter 4, I have used PduB' protein and dissected it into its individual domains from it to get domain1 and domain2 separately. They were compared to native single-domain protein PduA and a structurally similar variant of PduA was developed to get insights into the significance and specificities of a structure and sequence. In Chapter 5, I have studied the implications of substituting the PduA protein *in vivo* with a structurally similar protein but harboring a different sequence. Also, the positional role of a hexamer and a pseudohexamer was attempted by gene swapping *in vivo* using PduA and PduBB'. The shell proteins are known to self-assemble naturally and form a molecular sheet in 2D. This intrinsic property makes it easy to use them as scaffolds and due to their build up from amino acids backbone, they can be used to synthesize gold nanoparticles over their surface. In Chapter 6, I utilized this property and used PduA and its mutant PduA[K26A] self-assembling into 2D and 3D architecture to develop nanocatalysts and study the impact of scaffold architecture in tuning the properties of gold nanoparticles.

CHAPTER – 2

Material and Methods

This Chapter comprises detailed information about the materials and methodology used to carry out all the research work for this thesis. The entire work includes molecular biology, biophysical, biochemical, and spectroscopic techniques.

2.1 Materials

Table 2.1 List of chemicals and reagents

Name of the chemical/reagent	**Company/ Source**
3-methyl-2-benzothiazoline hydrazine	Sigma Aldrich, India
3,3′,5,5′-Tetramethylbenzidine	Sigma Aldrich, India
Acrylamide	Hi-Media, India
Ado-Cobalamin	Sigma Aldrich, India
Ammonium persulphate	Sigma Aldrich, India
Auric chloride	Sigma Aldrich, India
B-PER II, Bacterial protein extraction reagent	ThermoFischer Scientific, India
Bis-acrylamide	Sigma Aldrich, India
Bradford reagent	Sigma Aldrich, India
Citric acid monohydrate	Himedia, India
Cyanocobalamin	Sigma Aldrich, India
DMSO	Sigma Aldrich, India
DNase	Sigma Aldrich, India
EDTA disodium salt dehydrate	Hi-Media, India
Ethanol	Hayman, UK
Glycine	Hi-Media, India
HCl	Molychem, India
HEPES	Hi-Media, India
Hydrogen peroxide	Sigma Aldrich, India
Imidazole	Hi-Media, India
IPTG	Hi-Media, India
LB Agar	Hi-Media, India
Luria Broth	Hi-Media, India

Lysozyme	Hi-Media, India
Magnesium chloride hexahydrate	Hi-Media, India
Magnesium sulphate heptahydrate	Hi-Media, India
Ni-NTA resin	Qiagen, Thermo Fischer Scientific, India
Nickel sulphate hexahydrate	Sigma Aldrich, India
4-Nitrophenol	Hi-Media, 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	Hi-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
Sypro Orange	Sigma Aldrich, India
TEMED	Sigma Aldrich, India
Tris Base	Hi-Media, India

Table 2.2 Major Instruments used

Name	**Company**
Bio-Rad Gel electrophoresis	BIO-RAD, USA
CFX96 real-time PCR	BIO-RAD, USA
Circular Dichroism Spectroscopy	JASCO, Japan
Dynamic light scattering	Malvern, UK
High-speed centrifuge	Eppendorf, Hamburg, Germany

Incubator	Eppendorf, Hamburg, Germany
Isothermal Titration calorimetry	MicroCal-iTC200, Malvern Pananalytical
Microtiter Multimode plate reader	Biotek-EonTM, Vermont, USA
Steady state fluorescence	Edinburgh Instruments
Transmission Electron Microscopy	JEOL, Japan
UV-Vis Spectrophotometer	Cary 60 UV , Double Beam Agilent Technology
X-ray Diffraction	Bruker ECO D8, Bruker, US

Table 2.3 Information about protein variants involved in this study

	Detailed information	**Reference**
***In vitro* dimer protein variants**		
PduB' [WT]	PduB' wild type	gift from Prof. TA Bobik (Iowa State University, Ames, USA)
PduT [WT]	PduT wild type	This study (generated in Sinha Lab)
PduA[D]	PduA dimer	This study (generated in Sinha Lab)
PduA[CP-D]	Circularly permuted PduA dimer	This study (generated in Sinha Lab)
***In vitro* single-domain protein variants**		
PduA [WT]	PduA wild type	gift from Prof. TA Bobik (Iowa State University, Ames, USA)
PduA [K26A]	PduA mutant	gift from Prof. TA Bobik (Iowa State University, Ames, USA)
PduB'[D1]	Dissected PduB' doamin1	This study (generated in Sinha Lab)
PduB'[D2]	Dissected PduB' doamin2	This study (generated in Sinha Lab)
PduB'[CP-D2]	Circularly permuted PduB' doamin2	This study (generated in Sinha Lab)

***In vivo* modified strains**		
Salmonella enterica serovar Typhimurium LT2 (BE293)	With pkD46 plasmid	gift from Prof. TA Bobik (Iowa State University, Ames, USA)
SSL1	pkD46 cured *Salmonella enterica* serovar Typhimurium LT2 (BE293)	This study (generated in Sinha Lab)
SSL2	ΔPduA::ARC	This study (generated in Sinha Lab)
SSL3	ΔPduA: PduB'[CP-D2]	This study (generated in Sinha Lab)
SSL4	ΔPduA:PduBB'::PduBB':PduA	This study (generated in Sinha Lab)

2.2 Methodology

2.2.1 MCP Purification

For Pdu MCP purification previously developed protocol was followed[1,2]. Briefly, wild-type *Salmonella enterica* LT2 and its operon-modified strains were cultured in non-carbon E medium (NCE), supplemented with 0.6 % 1,2-PD, 0.5 % succinic acid and 1mM of MgSO4 and incubated at 37^{o}C for 16 h so the OD_{600} ~ 1-1.2 is reached. The cells were harvested, washed twice and resuspended in Buffer- A (50 mM Tris Base (pH 8), 500 mM KCl, 25 mM NaCl, 12.5 mM MgCl2, 1.5 % 1,2-PD) containing 75% bacterial protein extraction reagent (BPER-II) with 2-3mg DNase, 0.4mM phenylmethane sulfonyl fluoride (PMSF), and 1 mg/ml of lysozyme. The re-suspended cells were kept at room temperature for 30 minutes on a shaker at 45-50 rpm, followed by 5min on ice. The lysate obtained was separated for supernatant (containing MCP) from debris by centrifuging at 12,000 X g for 5 min, 4 ^{o}C. MCPs obtained in the supernatant were further centrifuged at 20000 X g for 20 min, 4 ^{o}C. The pellet now contains the MCP which was re-suspended in Buffer A containing 60% of B-PER II and 0.4 mM PMSF. It was centrifuged at 20,000 X g for 20 min at 4 ^{o}C and the thin film obtained was further re-suspended in pre-chilled Buffer B (50 mM Tris Base pH 8, 50 mM KCl, 5 mM MgCl2, 1 % 1,2-PD). Further, it was

centrifuged at 12,000 X g for 5 min at 4 ^{o}C and the supernatant obtained here contained PduMCP which was collected and stored at 4 ^{o}C.

2.2.2 Overexpression and purification of proteins

2.2.2.1 Expression and purification of scaffold proteins

The shell protein PduA cloned into pET-21a vector and its mutant PduA[K26A] into pET-41a vector with 6xHis-tag at N-terminal were transformed in BL21-DE3 expression host cells and streaked on an antibiotic resistance agar plate. Initially, both constructs were expressed in the pET-41a vector. PduA protein has less solubility than PduA[K26A] mutant. To increase the yield and pursue further studies, PduA construct was cloned in the pET-21a vector. A single colony from the streaked plate was inoculated for primary culture in 10mL of LB media with ampicillin (100µg/mL) for PduA and kanamycin (50µg/mL) for K26A mutant and incubated at 37 °C, 180rpm for 12 h. 1% of the overnight grown culture was used for secondary culture in 600ml of LB media. The culture was then incubated for 2-2.5 h at 37 ^{o}C, 180 rpm till OD_{600} reached 0.6-0.8 and induced with 0.5mM IPTG for 6h at 28 °C. The cells are harvested post-induction by centrifugation at 6000x g for 10min and lysed in column buffer containing 50mM Tris-HCl, 200mM NaCl, 5 mM Imidazole (pH 7.5) by subjecting to sonication. The clear supernatant was obtained after separating the cell debris by centrifugation at 11000x g for 30 minutes and loaded onto a Ni-NTA (Nickel-nitrilotriacetic acid) column pre-equilibrated with column buffer. The column was then washed with column buffer containing 50 mM imidazole and the protein was eluted with 200 mM imidazole. The purity of fractions was evaluated by SDS-PAGE (Sodium dodecyl sulfate-polyacrylamide gel electrophoresis) and protein concentration was determined using Bradford reagent.

2.2.2.2 Expression and purification of protein fold variants

Protein fold variants were expressed from IPTG inducible pET21-plasmids by transforming in BL21 (DE3) PlysS expression host strain using a similar protocol used for wild-type shell proteins [1,3,4]. A single colony was inoculated in 10ml LB broth at 37 ^{o}C, 180rpm, 12 h for primary culture

with ampicillin (100µg/mL), and chloramphenicol (34 µg/mL) antibiotics for all synthetic dimer and single domain variants and PduT [WT]as well, PduA [WT] with ampicillin (100µg/mL), and PduB'[WT] with kanamycin (50µg/mL). 1% of primary culture was used to scale up culture volume to 800ml LB for overexpression of proteins. The cultures were induced at 0.5mM IPTG at 28 ^{O}C, 180 rpm for 8 h to mid-exponential phase (OD_{600} of 0.6-0.8). Post induction the cells were harvested by centrifugation at 6000g for 10min and lysed by sonication in a column buffer containing 50mM Tris, 200mM NaCl, and 5mM Imidazole (pH 7.5). For purification, Ni-NTA (Nickel-nitrilotriacetic acid) affinity chromatography methodology was used as previously described for wild-type shell proteins. The supernatant obtained after separating the debris from lysate was loaded onto pre-equilibrated Ni-NTA resin with column buffer. The column was washed with a buffer containing 50 mM Imidazole. The proteins were eluted with 200mM Imidazole, dialyzed, in 10mM sodium phosphate buffer (pH 7.4) and the concentration of proteins was estimated using the Bradford assay.

2.2.3 Gold nanoparticle fabrication on shell proteins

The shell proteins PduA and PduA[K26A] were used at a concentration of 10µM for the fabrication of nanoparticles and incubated with 500µM $AuCl_3$. Here, $AuCl_3$ interacted with the surface residues of the shell proteins and was reduced to Au (0) from Au (III). The gold nanoparticle fabrication was monitored by measuring the reduction kinetics[5,6] by UV-Vis Spectroscopy using Carry 5000 UV-Vis Spectrophotometer for 90 min with 1 min time interval while scanning the spectra from 250nm to 900nm wavelength.

2.2.4 Catalysis of Shell Protein Gold Nanoparticles

2.2.4.1 Catalytic Activity of Shell Protein Gold Nanoparticles

The stability and catalytic activity of fabricated gold nanoparticles on shell protein surface were studied by following a standard reaction where p-nitrophenol is reduced to p-aminophenol in the presence of $NaBH_4$ as a reducing agent[84,85] using UV-Vis Spectroscopy (Carry 5000 UV-Vis Spectrophotometer). 50µl of 1mM p-nitrophenol was diluted in 900µl Milli-Q water and mixed

with 10µl of 100mM $NaBH_4$ in a quartz cuvette. 20µl of shell-AuNPs were added before spectral reading from 250-600nm wavelength. The reaction was monitored for 30min with a 1-minute time interval.

2.2.4.2 Peroxidase-like catalytic activity of Shell Protein Gold Nanoparticles

The peroxidase-like catalytic activity of freshly prepared shell protein gold nanoparticles was determined by measuring the formation of a blue charge-transfer complex of diamine from TMB at 652nm (ε=39000$M^{-1}cm^{-1}$)[9]. The peroxidase activity measurement was done in 96 well plate in a 200µl reaction. 20µl shell protein gold nanoparticle composite was mixed with different concentrations of TMB. The kinetics reaction was monitored for 30 minutes by measuring absorbance at 652nm using a microtiter multimode plate reader (Biotek – EonTM). At the end of the reaction, the spectra were recorded from 500-800nm. Similarly, temperature-dependent activities were performed where nanocomposites were incubated for 20min at a particular temperature (20-80 ^{O}C), and respective pH activity was performed in glycine buffer (pH 3-9).

2.2.5 Design and generation of synthetic protein variants

Dimer and single-domain synthetic protein fold variants were designed by considering protein structures and exploring the domain sequence. Dimer protein variants were generated by using a sequence from the PduA[WT] protein. The native linker (GDVYGNEA) from PduB'[WT] was used to fuse the single fragments. For single-domain studies, PduB'[WT] dimer was dissected into two independent domains from the linker region joining two fragments. For circular permutation, respective domain structures from two proteins (PduA, PDB ID-3NGK; PduB', modeled from PDB ID-4FAY) were aligned using PyMol software[10], and differences in N and C terminals and linker regions were identified. The structures were then modified as per the desired structure by fusing old N and C terminals and generating new terminals using domain sequences.

2.2.6 Cloning

Plasmid DNA constructs for all synthetic protein variants and PduT[WT] were cloned in the pET21a vector between Nde1 and Xho1 restriction sites. PduA[WT] and PduB'[WT] constructs were already present in the lab. For cloning, designed gene fragments were amplified using polymerase chain reaction (PCR) and cloned into linearized vector through double digestion using the Gibson cloning method[11]. Positive clones were screened on an ampicillin antibiotic agar plate (100 μg/mL) and identified using colony PCR. Further, they were verified using Sanger sequencing.

2.2.7 Structure prediction, structure alignment, and domains flexibility

All the three-dimensional structures of synthetic dimer and single-domain protein variants were generated using AlphaFold2 implemented in ColabFold[12,13]. Designed sequences involved in the present study were given as inputs to generate structures. For visualizing structures, structure alignments, and RMSD calculations, the PyMOL software was used. Domain flexibility was determined using CABS-flex 2.0 [14] and MEDUSA[15] web servers.

2.2.8 Methodology to develop the operon-modified strain using CRISPR-Cas9 gene-editing tool

For the development of operon modified strain, firstly, lab strain *Salmonella enterica* serovar Typhimurium LT2 (BE293) which carries pkD46 plasmid was cured. It was cured by repeatedly growing the strain in the absence of the antibiotic, till it gained the sensitivity. This cured mutant (SSL1) was employed for genome editing studies using CRISPR Cas9 Technology[16]. Initially, plasmid pCasM was electroporated into *Salmonella enteria* LT2 (SSL1) which contains both Cas9 and a λ-Red recombineering system. Electroporation was performed in a 2-mm cuvette at 2.5 kV, 200 Ω, and immediately 1ml ice-cold SOC media was added to resuspend the cells and incubated at 37^{o}C. The cells were allowed to recover for 2 h at 37^{o}C. The transformants were then selected on a Kanamycin antibiotic selection plate (50μg/ml) and used to prepare electrocompetent cells for the insertion of an antibiotic resistance cassette. The electrocompetent cells were prepared at ~0.6 $OD_{600.}$ The harvested cells were then washed three times and finally resuspended in ice-chilled water containing 10% glycerol. The cells were concentrated to 50μl volume and used for electroporation. The antibiotic resistance cassette, used here, the ampicillin resistance cassette

(ARC) was prepared by Overlap Extension PCR to be inserted at the target site. The ampicillin resistance gene (*bla*) was PCR amplified and fused with upstream and downstream homologous arms of the target site. The generated ARC cassette was electroporated into the above prepared electrocompetent cells (50µl). Transformants were screened using an agar medium containing Kanamycin (50µg/ml) and Ampicillin (100µg/ml) antibiotics. The colonies were screened by PCR screening method and sequencing after genomic DNA isolation. Positive clones were then further used to prepare electrocompetent cells to transform the desired gene (donor DNA). This mutant (SSL2) is referred to as PduA[WT] substituted strain (with ARC). The desired gene to be inserted was also generated by the Overlap Extension PCR method containing the upstream and downstream homology arms. Donor DNA and psgRNA were electroporated to the above-prepared electrocompetent cells respectively. Transformants were screened on Kanamycin (50µg/ml) and spectinomycin (200µg/ml) antibiotics. Further positive modified clones were confirmed using PCR screening method and Sanger sequencing. The positive mutant obtained is referred to as SSL3.

2.2.9 Growth assay

For growth assay studies with wild-type *Salmonella enterica* LT2 and its operon-modified strain, 5ml overnight grown culture in LB media was used to inoculate 10ml LB supplemented with 0.6% 1,2-PD. The culture was allowed to grow for 6 h and then cells were harvested and washed three times with 1x NCE medium containing 0.6% 1,2-PD and 1 mM MgS04. The washed cells were suspended to an OD600 of 0.15 in 1x NCE medium (supplemented with 0.6% 1,2-PD; 0.3 mM each of valine, isoleucine, leucine, and threonine; 150 nM CN-B_{12}) and allowed to grow at 37°C with constant shaking. For growth analysis of both strains, 200 µl sample was collected at intervals and optical density was measured at 600nm. The growth curves were repeated at least three times in duplicate and representative curves are shown.

2.2.10 Characterization techniques

2.2.10.1 Transmission electron microscopy (TEM)

TEM images of fabricated gold nanoparticles on the protein surface and protein variants were obtained by negative staining using JEM 2100 TEM (JEOL, USA) operated at a voltage of 120 kV. The samples were added to fomvar coated grids (300 mesh size). The grids of protein samples were prepared by adding 10µl of 10µM sample followed by staining with 1% uranyl acetate (freshly prepared and filtered) and washing with Mili-Q (pH 7.3-7.5). For fabricated gold nanoparticle-protein conjugates, 10µl of the conjugated sample was added directly to the grid. The excess sample was wicked off, followed by washing with Mili-Q (pH 7.3–7.5) and drying before visualization.

2.2.10.2 X-Ray Diffraction Studies

X-ray diffraction (XRD) was employed to evaluate the crystallinity of the fabricated nanoparticles. XRD was recorded using Bruker ECO D8 advance with Cu Kα radiation wavelength of 0.1542nm at a voltage of 40 kV and a current of 25 mA. The data was collected from 10 to 80 two theta (2θ).

2.2.10.3 Isothermal Titration Calorimetry

ITC experiments were performed on a MicroCal PEAQ ITC200 system (Malvern) and carried out at physiological pH 7.4 and room temperature 25 ^{O}C by titrating 10µM protein solution into a sample cell containing 500µM gold precursor solution. Titration involved 39 injections (1µl per injection) of proteins at 150s intervals into a sample cell (volume 280ul) with continuous stirring at 750rpm. The heat profile of dilution in the buffer was subtracted from titration data and analyzed to determine all binding thermodynamic parameters.

2.2.10.4 Circular Dichroism Spectroscopy

To analyze the secondary structure of synthetic protein variants and the effect of fabrication of gold nanoparticles on the structure of proteins secondary structure estimation was carried out using a CD spectrophotometer (Jasco J-1500, CD spectrophotometer). CD spectra of wild-type, synthetic protein variants, and shell protein gold nanoparticle conjugates were recorded by taking 200 µl protein sample (12µM) using a quartz cell with a path length of 0.1 cm. The samples with

fabricated nanoparticles were centrifuged at 13000 rpm for 30 min and the pellet was re-suspended to record the CD. The CD spectra of the supernatant were also recorded. The spectra were recorded from 260 to 195nm with an accumulation of three scans and a scanning speed of 200nm/min. The sample temperature was maintained at 25 °C using a mini circulation water bath (Jasco MCB-100) connected to the water-jacketed sample holder chamber. To check the effect of temperature on secondary structures of protein variants CD spectra were recorded from 25 °C to 95°C at the rate of 2.5 °C per min. For post-melt studies that temperature was gradually allowed to cool down to 25 °C and then the spectra were recorded.

2.2.10.5 Dynamic Light Scattering

Dynamic Light Scattering **(**DLS) studies for the size distribution of protein variants were performed using a Zetasizer Nano ZSP (Malvern Instrument, UK). The scattered intensity was measured at a backscattered angle of 173°. For all measurements, 200 µl of 12µM protein samples were taken in a disposable polystyrene cuvette and the intensity percentage was recorded. All measurements were done in triplicates. For temperature-dependent size variation of the protein variants glass cuvette with a 1ml protein sample was used and the number percentage was plotted with respect to the temperature to get the maximum population and avoid aggregates. All measurements were done in triplicates.

2.2.10.6 Fluorescence Spectroscopy

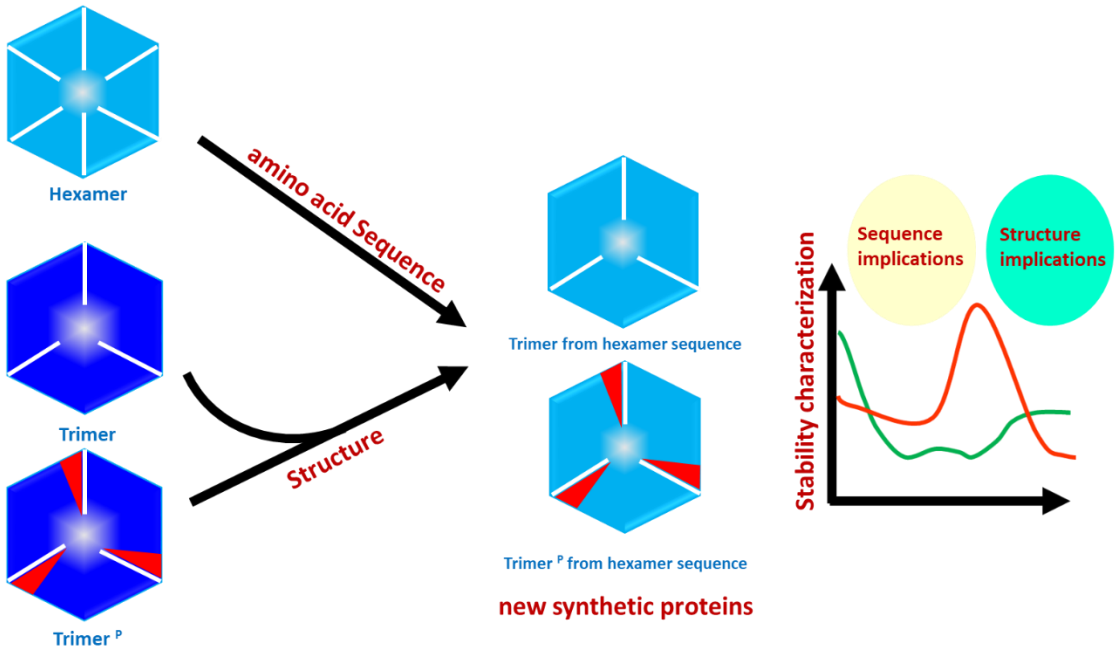
A fluorescence spectrofluorometer (FS5, Module SC-5, Standard cuvette holder from Edinburgh Instruments, UK) was used to study the formation of gold nanoclusters. The shell protein-gold nanoparticle conjugate was excited at 370nm and the emission was collected from 500nm to 700nm with excitation and emission bandwidth of 5nm each. The emission spectra were recorded with a 1-minute interval time. Intrinsic steady-state fluorescence measurements for protein variants were carried out using 400 µl of 12µM protein samples and were excited at 285 nm and emission spectra were recorded from 300 to 400 nm with a bandwidth of 3 nm.

2.2.10.7 Differential scanning fluorimetry

DSF studies were carried out using a CFX96 real-time PCR system (Bio-Rad) equipped with a 96-well block. Each protein variant at 12 μM concentration was mixed with 20x SYPRO Orange dye (Sigma) and was loaded into 96-well thin-wall PCR plates (BioRad). Plated were sealed with optical-quality sealing tape (BioRad). The temperature was increased from 20 °C to 95 °C in 1 °C/min at each cycle increments with an equilibration time of 1:30min at each temperature. The change in SO fluorescence is monitored over time (λex = 450−490 nm, λem = 610−650 nm) with the help of the CFX96 Touch System, which includes one channel with an LED-filter photodiode combination designated for single-color fluorescence resonance energy transfer (FRET). The transitions measured on the FRET channel were displayed as relative fluorescence units (RFU) with respect to temperature. Tm values were calculated from raw DSF curves after baseline subtraction and multiple peak fitting.

CHAPTER – 3

Potential role of domain dimerization in Bacterial microcompartment



3.1 Introduction

The first chapter has introduced bacterial microcompartments (BMCs) with detailed information about these unique protein organelles which are conditionally expressed and carry out specialized metabolic processes in diverse bacteria[3,19,33]. Their outer envelope is composed of only proteins which encase enzymes within, and are involved in specific cascading reactions. The protein envelope is built from a few thousand protein subunits belonging to bacterial microcompartment shell proteins domain family (Pfam00936)[1,37]. Shell proteins are seen as structurally equivalent proteins with flat hexagonal symmetry which forms the basis of the assembly of facets of BMC icosahedral structure. To date, we have a good understanding of the architectural principles of BMC shell proteins[1,34,36,45,94–96]. However, the complexity of a shell with more than one protein type is still a puzzle to solve. The shell proteins attain two different types of oligomerization, viz., hexamer and psuedohexamer. Hexamer is formed from self-assembly of six canonical BMC domains belonging to the BMC domain family. Another variation as psuedohexamers, is a result of gene duplication events where two BMC domains are fused in a tandem manner to form a dimer and three units of it assemble to form a trimeric pseudohexamer[50,97,98]. The overall shape of a pseudohexamer resembles a canonical hexamer. Furthermore, increasing the complexity, some shell proteins are also structurally correlated through circulation permutations[32,99]. Circular permutations result in reordering of amino acid sequences but retain similar topology during evolution and play important role in protein folding and function[100–102]. Intentional circular permutations have been introduced to engineer proteins for altering activity, improving stability, reducing proteolytic susceptibility and controlling regulation[103–108]. The BMC domain fold consists of three α helices and four β strands[1,99]. Permuted BMC fold has a secondary structure element occurring in different order than canonical BMC fold but with similar 3D topology retained. The secondary structure elements contributing to C-terminal region in canonical fold, contributes to the N-terminal region in circular permuted BMC fold. Structural studies have revealed circularly permutant variants of a single canonical protein fold forming hexamer and also in protein folds existing as tandem repeats forming psuedohexamers[41,47,48,95]. The role of existence of a particular amino acid sequence and a distinct topology is lesser understood in bacterial microcompartments. In this chapter, the potential role of domain dimerization and circular permutation is addressed by exploring one of the complex metabolosomes known, 1,2 propanediol

microcompartment (1,2 MCP) with eight different shell proteins. To get insights into the puzzle of structural diversity and sequence variation, native sequence of one of the major and well-studied shell protein, PduA[42,59,61,65] is explored and generated synthetic dimer-protein fold variants. Two naturally occurring dimer-domain proteins were considered for this study. PduBB' is one dimer-domain protein oligomerizing as a pseudohexamer, critical in MCP assembly and is expressed from overlapping genes[61,65]. It is expressed from an overlapping gene as two proteins namely, PduB and PduB' with identical sequences except the extended N-terminal region in PduBB'. PduB is translated from the first start site comprising N-terminal 37 amino acids and the remaining 233 amino acids. PduB' on the other hand is a short-length protein that excludes the initial 37 amino acids and is translated from the second start site having 233 amino acids. PduBB' is a well-studied protein for the role of N terminal 37 amino acids where it is emphasized that the N-terminal segment is a disordered region and involved in binding the outer shell to its enzymatic core[72,73]. PduT is another example of a dimer shell protein which is a pseudohexamer. PduT is not an essential protein for bacterial microcompartment formation but it contains a 4Fe-4S cluster at the center of the pore and may have significant involvement in an electron transfer process[42,71,98]. However, the understanding and significance of different oligomerization states as psuedohexamers and hexamer proteins is not clear. So, two new dimer protein variants were generated which are structurally similar to pseudohexamer proteins i.e., PduB' and PduT using the PduA sequence which forms hexamer. One is a circularly permutant dimer which is structurally similar to PduB'. Another, a non-permuted dimer whose topology is equivalent to naturally existing PduT and hence compared to PduT shell protein.

3.2 Generation strategy of synthetic dimer-protein variants

This is interesting that the shell proteins belonging to the same domain family undergo duplication events during the course of evolution and result in different oligomerizations. The question here remains why MCP needs such diversity and complexity? Each specific protein sequence dictates a specific conformational topology and functional role. Here, we addressed a question, whether a particular sequence from one shell protein (single-domain), generating a different topology

Protein variants	Amino acid sequence
PduA [WT]	MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDAGAAAARNV GEVKAVHVIPRPHTDVEKILPKGISQ
PduA [D]	MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDAGAAAARNV GEVKAVHVIPRPHTDVEKILPKGISQGDVYGNEAMQQEALGMVETKGLTAAIEAADAMVKSANVMLVGY EKIGSGLVTVIVRGDVGAVKAATDAGAAAARNVGEVKAVHVIPRPHTDVEKILPKGISQ
PduA [CP_D]	EVKAVHVIPRPHTDVEKILPKGISQMQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVI VRGDVGAVKAATDAGAAAARNVGGDVYGNEAEVKAVHVIPRPHTDVEKILPKGISQMQQEALGMVETK GLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDAGAAAARNVG
PduT [WT]	MSQAIGILELTSIAKGMELGDAMLKSANVDLLVSKTICPGKFLLMLGGDIGAIQQAIETGTSQAGEMLVDSLV LANIHPSVLPAISGLNSVDKRQAVGIVETWSVAACISAADRAVKGSNVTLVRVHMAFGIGGKCYMVVAGD VSDVNNAVTVASESAGEKGLLVYRSVIPRPHEAMWRQMVEG
PduB' [WT]	MAEKSCSLTEFVGTAIGDTLGLVIANVDTALLDAMKLEKRYRSIGILGARTGAGPHIMAADEAVKATNTEVVS IELPRDTKGGAGHGSLIILGGNDVSDVKRGIEVALKELDRTFGDVYGNEAGHIELQYTARASYALEKAFGAPIG RACGIIVGAPASVGVLMADTALKSANVEVVAYSSPAHGTSFSNEAILVISGDSGAVRQAVTSAREIGKTVLAT LGSEPKNDRPSYI

Figure 3.1 Protein sequences of the desired dimer-domain constructs for the present study. PduA [WT] (top most), PduA [D] (second from the top), PduA [CP_D] (middle), PduT [WT] (second from the bottom), PduB' [WT] (bottom). Linker region is highlighted in green color in between the sequence.

equivalent to its native counterparts(dimer-domains), will give rise to the same biophysical properties as their native protein sequence and topology. For this purpose, we used PduA sequence [Figure 3.1] and the linker region sequence present in PduB' shell protein to generate a synthetic tandem bacterial microcompartment domain proteins. For PduA dimer mentioned further as PduA [D], direct two PduA domains were fused by a linker sequence (GDVYGNEA) from PduB' and a circularly permutant dimer of PduA mentioned further as PduA [CP_D] was generated by using a circularly permuted PduA domain sequence and fusing two such domains using the same linker sequence [Figure 3.2 (c)]. To generate a circularly permutant domain of PduA [WT] adapting topology similar to PduB'[WT], the two protein's monomers were aligned using the PyMOL software[87]. On alignment we observed that PduA[WT] (PDB 3NGK) overlapped with the second domain of PduB'[WT] (modelled from PDB 4FAY) with an RMSD score of 0.85Å than domain 1 having RMSD of 1.49 Å. It was seen that they have a similar topology but have different N and C terminals highlighting a circular permutation. So, the existing N and C terminal of PduA[WT] sequence were joined by a linker and new termini were created at a site similar to PduB'[WT] protein structure (domain 2 fold) [Figure 3.2 (a)]. Once we designed the sequences for both variant types, the structures were generated using the AlphaFold2[89,90] and visualized by PyMOL. The multiple sequence alignment of PduA [WT] with PduT [WT] showed 30.8% sequence similarity and 18% with PduB' [WT] [Figure 3.2 (b)]. It suggests that PduA[WT] is more related to PduT[WT] than PduB'[WT]. PduT[WT] and PduB'[WT] have only 16% sequence similarity. The new protein variant sequences when aligned to PduB'[WT], the similarity percentage increased to 20% with PduA [CP-D] and remained 17.8% with PduA [D]. On aligning with PduT the similarity percentage decreased to 25% with PduA [CP-D] and 27.7% with PduA[D]. Thus, PduA[CP_D] generated after circular permutation has increased sequence similarity with PduB'[WT] which has undergone sequence rearrangement and now has a similar topology, through circular permutation of PduA domain. This points to the evolutionary significance of BMC domain variations.

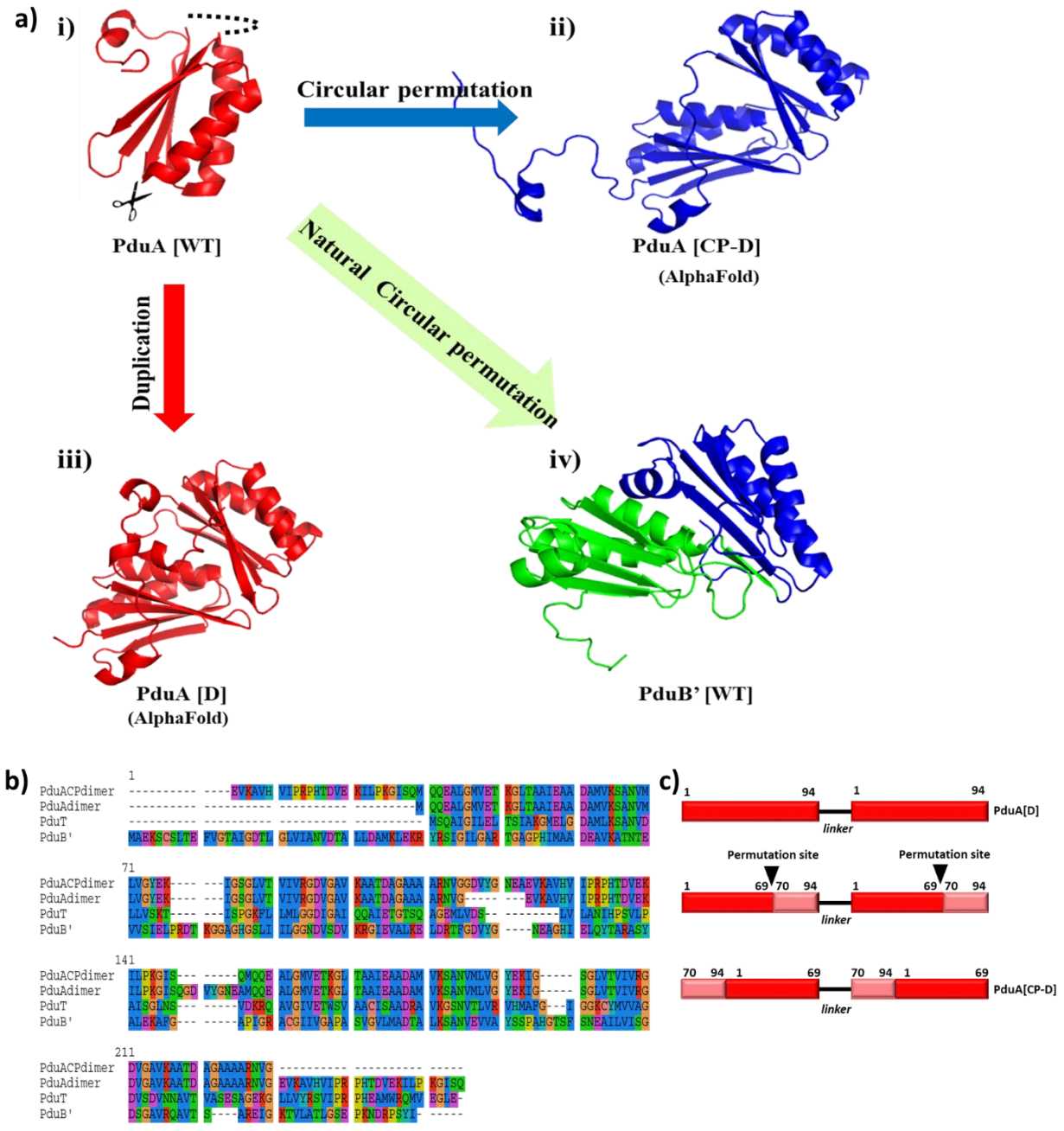


Figure 3.2 Schematics of correlation and generation of synthetic dimer-protein variants from WT proteins. a) AlphaFold2 was used to generate the structure of dimer variants and visualized using PyMOL to show the domain fold of protein variants. PduA[WT] structure highlights the cutting and pasting sites for circular permutation. b) Multiple sequence alignment of protein variants with PduA domain sequence using *ClustalW* c) Image showing

the design of dimer-protein variants from PduA[WT] sequence, PduA[D] (red), PduA [CP-D] (red and pink to distinguish circular permutation region). For circular permuted variants, the permuted site is highlighted with black arrow along with residue numbers.

3.3 Cloning overexpression and purification of synthetic dimer-protein variants

We used the Gibson cloning[88] methodology to obtain the synthetic dimer-protein variant constructs, PduA [D] and PduA [CP-D] and PduT [WT] construct. The polymerase chain reaction (PCR) was used to amplify the designed fragments with overlap regions corresponding to the fragments and the vector. The amplified products were then ligated with the linearized vector by following the Gibson procedure and transformed in a cloning expression strain DH5-α [Figure 3.3 (a)]. The positive clones were screened on an antibiotic agar plate and further verified using colony PCR and sanger sequencing [Figure 3.3 (b)]. Post verification of the positive clone it was transformed in a protein expression strain BL21 (DE3) PlysS cells for protein induction and purification. PduT[WT], PduA [D] and PduA [CP-D] proteins were overexpressed and purified using the same protocol used for other wild-type shell proteins[44,80,109]. The overexpression of proteins was induced at 0.5mM IPTG at 28^{O}C for 8 hr at an OD_{600} of 0.6. The cells were harvested and lysed post induction and checked for the induction of proteins. We found the proteins to be expressed in both soluble fraction and inclusion bodies. Although, the expression in a soluble fraction was less and most of the protein was present as inclusion bodies. We proceeded with the protein present in a soluble fraction as it was sufficient for our biophysical studies. Interestingly, we observed PduA [CP-D] to be more soluble like PduB'[WT] as compared to PduA [WT] and PduA [D]. For purification Ni-NTA affinity chromatography methodology was used as previously used for WT shell proteins. The N-terminal 6xHis-tag proteins were passed through the pre-equilibrated with Ni-NTA resin. The beads were equilibrated with column buffer which has a composition of 50mM Tris, 200mM NaCl, and 5mM Imidazole, pH 7.5. The proteins were eluted with same column buffer but at 200mM Imidazole concentration, pH 7.5. The concentration of proteins was estimated using the Bradford assay.

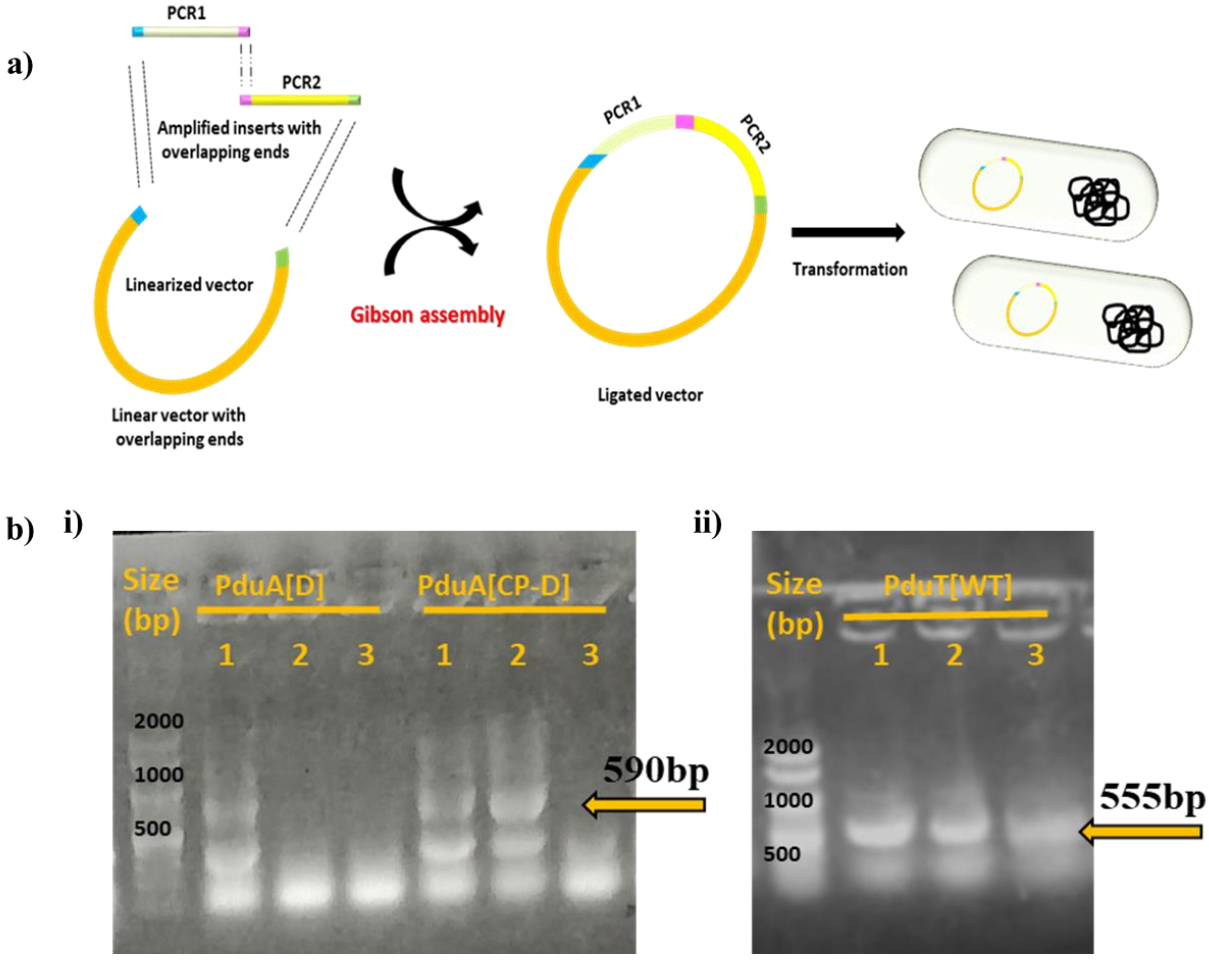


Figure 3.3 Schematics of Gibson Assembly protocol and corresponding gel images of PCR products. a) Schematic showing Gibson cloning method used to clone synthetic dimer protein variants. b) Agarose gel image of colony PCR products screening for positive clones (i) PduA[D] and PduA[CP-D]; (ii) PduT[WT]. Highlight bands are specific to PCR product respective to the desired insert size.

3.4 Biophysical studies to characterize the protein fold variants

To understand the folding, self-assembling, morphology and stability properties of the synthetic protein variants and compare them to their WT protein counterparts having similar structural topologies, we used an array of biophysical techniques.

3.4.1 Circular dichroism and steady-state fluorescence spectroscopy to estimate the secondary structure and folding of synthetic dimer-protein variants

3.4.1.1 Circular dichroism

The native shell proteins give an α-helix signature when studied for secondary structure using Far-UV circular dichroism (far-UV CD). In Far-UV CD spectra, an α-helix shows two minima at 222nm and 208nm. The CD spectra showed a presence of an α-helix secondary structure for new protein variants and hence indicated that the proteins are properly folded. PduT [WT] shell protein was also verified for secondary structure as it was cloned for these studies and was found to have an α- helix signature as evident from the spectrum [Figure 3.4 (a)]. However, closely monitoring the spectra it is observed that the spectrum does not exactly overlap for all the protein variants. Instead, each respective protein spectrum has its own distinct features. PduB'[WT] and PduT [WT] protein's CD spectra overlap with minima at 208nm and 222nm [Figure 3.4 (a)]. PduA and PduA[D] CD spectra show similar spectrum features whereas PduA [CP-D] has a distinct spectrum with more negative ellipticity at 208nm.

3.4.1.2 Steady-state fluorescence spectroscopy

Steady- state fluorescence spectroscopy measures the intrinsic fluorescence of aromatic amino acids present in the proteins. Table 3.1 gives information about the aromatic amino acids that are present with respect to the shell proteins and their synthetic protein variants. The WT proteins and proteins generated for this study have presence of either of combination of three different aromatic amino acids (tyrosine, tryptophan, phenylalanine). PduA[WT] has single tyrosine present and on excitation at λ_{ex} 285nm, it gives emission fluorescence wavelength at λ_{em} 312nm. PduB' [WT] has six tyrosine and four phenylalanine intrinsic fluorophores. When excited it gives fluorescence emission at λ_{em} 300nm. A similar emission λ_{em} 300nm was observed in PduA[CP-D] which has a

structurally equivalent topology and three tyrosine residues. Another interesting observation is with PduT [WT] protein containing all the three intrinsic fluorophores each two in number. It gives a fluorescence emission at λ_{em} 335nm. PduA [D] which has a topologically similar protein to PduT [WT] gave fluorescence emission at λ_{em} 328nm. [Figure 3.4 (b)]. The WT and the synthetic dimer proteins has their own characteristic fluorescence emission spectra. This further proves that the synthesized proteins are properly folded and different packing of protein structure gave different emission wavelengths.

Table 3.1 Table with information about the abundance of aromatic amino acid residues present in shell proteins and their synthetic dimer-domain protein variants used in this chapter.

	Name of aromatic amino acids present	Number of aromatic amino acids present
PduA WT	Tyrosine Tryptophan Phenylalanine	1 0 0
PduA dimer	Tyrosine Tryptophan Phenylalanine	3 0 0
PduA dimer_CP	Tyrosine Tryptophan Phenylalanine	3 0 0
PduT WT	Tyrosine Tryptophan Phenylalanine	2 2 2
PduB' WT	Tyrosine Tryptophan Phenylalanine	6 0 4

3.4.2 Dynamic light scattering and TEM analysis to demonstrate the ability of self-assembling of synthetic dimer-protein variants into higher order structures

Bacterial microcompartment domain proteins has intrinsic property of self-assembly. So, it was interesting to know if these protein domain variants have still retained intrinsic property of self-

assembly to form higher order structures. To address this, we performed dynamic light scattering (DLS) experiments and plotted the intensity percentage. We observed the broad size distribution of PduA [D] in size range of 50nm-100nm and 200nm-500nm whereas for PduA [CP-D] the size lies in the range of 50nm -200nm and another population in 200nm-700nm size range. These broad size distributions are in a similar range to WT proteins. PduA [WT] showed bigger size distribution of sheets from 200nm to 900nm and PduB' [WT] 50-200nm, 200-800nm as already reported[80,109]. PduT [WT] protein also showed size distribution from 50-150nm, 300-700nm. This suggests that protein variants can form higher order structures. Transmission electron microscopy (TEM) further showed that the protein variants retain self-assembling properties to form higher order supramolecular structures. PduA [WT] and PduB' [WT] protein when overexpressed and purified form 2D sheets[45,79,110]. Synthetic protein variants from PduA domain sequence when expressed and purified they resulted in interesting self-assembling structures. PduA [D] protein generated by fusing two domains of PduA [WT] forms tubes [Figure 4 (d)]. This could be due to the fact that on dimerization the angle between the consecutive hexamers is a little bent which resulted in tubes[30,69,111]. On the other hand, PduA [CP-D] forms long thin sheets similar to PduA [WT] [Figure 3.4 (f)]. PduB' [WT] as earlier reported forms extended 2D sheets [Figure 3.4 (g)]. PduT [WT] protein which has a tandem domain repeats and has a similar topology like PduA [D] is observed to form self-assembled extended sheets [Figure 3.4 (e)]. This result was very encouraging where we observe that protein variants can self-assemble to from higher order structures.

3.4.3 Native PAGE shows oligomerization and stability

We performed native page to further investigate the oligomerization state of proteins. Native PAGE separate the proteins according to their molecular mass and under native conditions. Therefore, the oligomerization state of the proteins remain preserved. However, the proteins were separated under three conditions: unboiled (UB), boiled for 10min (B) and in the presence of SDS (SDS). Unboiled samples retain the higher order assemblies, whereas boiling the samples will allow the disruption of higher order assemblies and denaturation of proteins. Under this treatment smallest possible oligomer or monomer state of proteins will also be separated. SDS treated samples will assay the stability of proteins towards detergents. 10% native page was used and samples were loaded as unboiled (UB), boiled for 10 min (B), and in the presence of SDS (SDS).

We observed that all proteins form multiple distinct higher-order assemblies [Figure 3.4 (c)] which are heat and SDS-resistant as evident in UB and SDS lanes.

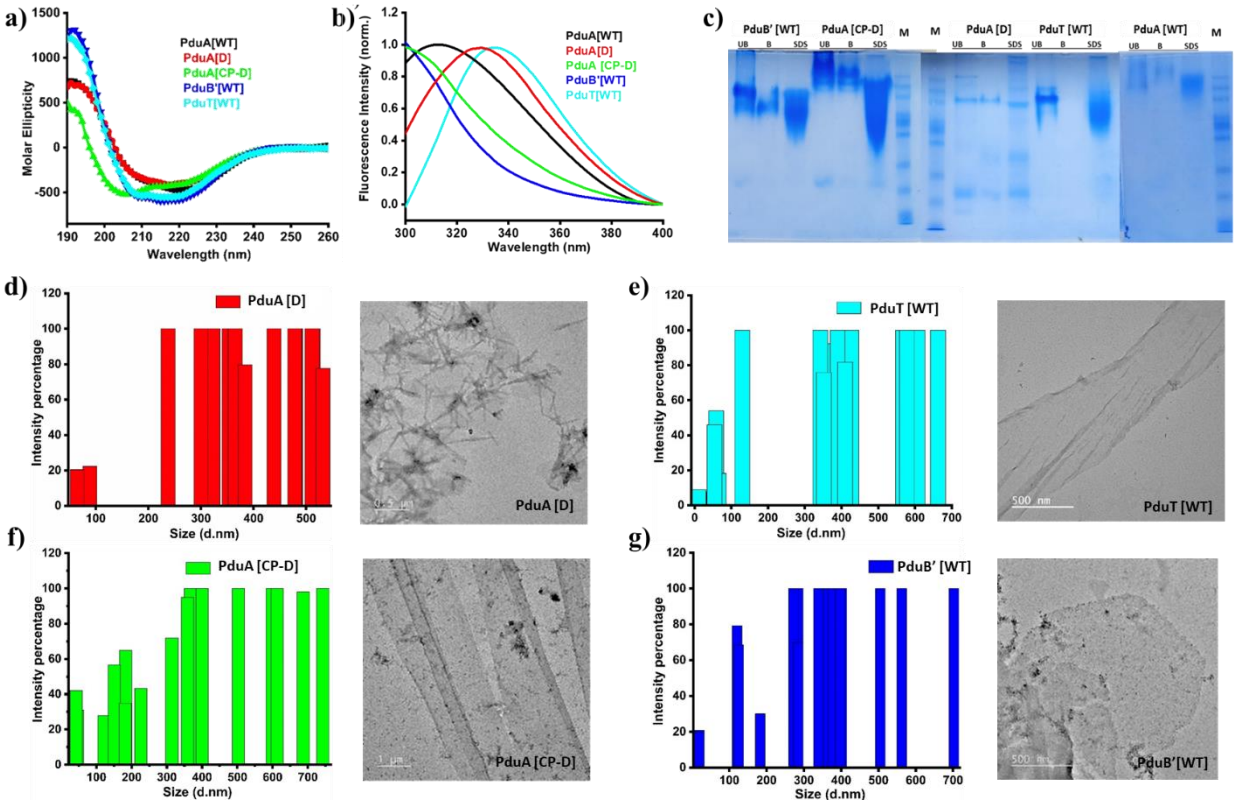


Figure 3.4 Characterization of WT and synthetic proteins using an array of biophysical techniques. a) Overlapped Circular dichroism spectra showing *α* helix secondary structure of PduA [WT] (black), PduA [D] (red), PduA [CP-D] (green), PduB' [WT] (blue) and PduT [WT] (cyan). b) Fluorescence emission spectra overlaid of PduA [WT] (black), PduA[D] (red), PduA [CP-D] (green), PduB' [WT] (blue) and PduT [WT] (cyan). c) Native-PAGE of PduA [WT], PduA [D], PduA [CP-D], PduB' [WT] and PduT [WT] under unboiled, boiled, and presence of SDS conditions. Dynamic light scattering (DLS) showing size distribution of proteins in the range of 100-800nm and TEM micrographs shows self-assembled structures of d) PduA[D] (red) size distribution (left) and TEM image(right). e) PduT [WT] (cyan) size distribution (left) and TEM image (right). f) PduA [CP-D] (green) size distribution (left) and TEM image (right). g) PduB' [WT] (blue) size distribution (left) and TEM image (right). All

protein variants assemble and form higher order structures. PduA [D] from tubes while other proteins self-assemble as extended sheets.

Till now we concluded that synthetic proteins are well folded and retain their self-assembling nature similar to WT proteins although the architecture is different. Next, we want to measure and compare how stable are these when subjected to high temperatures. We have performed temperature dependent stability studies using CD, DLS and DSF. CD gives an information about how stable is secondary structure elements w.r.t to temperature. Temperature dependent DLS will indicate the stability of higher order structure present in these proteins. Fluorescence based temperature ramp using DSF will provide an insightful information about the exposed hydrophobic patches on the sheet as well as the hydrophobic surfaces present in the core of the protein.

3.5 Temperature dependent stability studies of the protein fold variants using CD

Using circular dichroism, we measured the temperature dependent stability of two major shell proteins, PduA[WT] and PduB'[WT]. These shell proteins have different oligomerization state, hexamer being exhibited by PduA[WT], oligomerized from a single-domain and pseudohexamer by PduB'[WT], oligomerized from a dimer-domain. CD thermal melt studies were performed from 25 ^{o}C to 95 ^{o}C with a rate of 2.5/min using 12µM of proteins. We observed that PduA[WT] doesn't show significant uncoiling of secondary structure when subjected to thermal melt while temperature ramping from 25 ^{o}C to 95 ^{o}C. This implies that the secondary structure is retained with strong interactions which does not break on heating. Whereas, PduB' [WT] shows unfolding after 80 ^{o}C where it loses its native conformation. The negative maxima at 208nm shifts to 200nm after 80 ^{o}C indicating uncoiling of secondary structure. This is contrasting behavior of PduB'[WT] pseudohexamer protein compared to PduA[WT]. The post melt spectra for both proteins show structures are retained after cooling from 95^{o}C to 25^{o}C as it overlaps with the pre-melt spectra. It suggests that PduA[WT] is held with more strong interactions and consecutively more compact as compared to PduB'[WT]. Therefore, the structure of PduB' [WT] is not rigid as PduA[WT] rather it shows flexibility. To determine the melting temperature of proteins, CD signal at 222nm was

monitored and fraction of unfolded protein was calculated and plotted as a function of temperature and fit to Boltzmann distribution (Eq.1).

$$T = \frac{A_1 - A_2}{1 + e^{(T - Tm)/dT}} \quad \text{- equation (1)}$$

(Where A_1 and A_2 are the initial and the final values, T_m is the midpoint of the transition, and dT is the slope of the transition)

It is seen that, PduA[WT] protein shows the absence of proper S-shaped curve when fraction of unfolded protein was plotted. On the other hand, pseudohexamer protein PduB'[WT] Boltzmann fit gave T_m of 81.4 ± 2.4 ^{o}C [Figure 3.5 (d, e)]. We wondered whether this characteristic is specific to pseudohexamer proteins. Therefore, we performed temperature dependent CD studies with other dimer protein and compared. PduT [WT] [Figure 3.5 (b)] temperature dependent CD spectra shows decrease in negative ellipticity at 80^{o}C, thus indicating loosening its native conformation which is retained after cooling to 25^{o}C. The Boltzmann fit gave the T_m of 83.4 ± 1.6 ^{o}C. On analyzing other synthetic dimer protein variants generated from PduA[WT] sequence i.e, PduA [D] gave a T_m of 89.01 ± 4.5 ^{o}C [Figure 3.5 (a)] where as PduA [CP-D] gave T_m 85.2 ± 9.3 ^{o}C [Figure 3.5(c)]. Here, the temperature dependent CD spectra of domain protein variants data illustrate that the proteins secondary structure elements are stable and more compact in the case of PduA [WT] and PduA[WT] sequence variants i.e., PduA [D], PduA [CP-D], with increasing temperatures. BMC native dimer proteins PduT[WT] and PduB'[WT] shows a decrease in CD intensity with increasing temperature which signifies the change in native conformation. Interestingly, PduA[WT] protein variants on dimerization showed similar temperature-dependent behavior as of PduA [WT] [Figure 3.5(e)]. This finding is surprising as in circularly permuted variant of PduA[WT], the topology is similar to PduB'[WT] and expects to behave similar to PduB'[WT]. The direct dimer protein variant of PduA[WT], having similar topology to PduT[WT] also doesn't unfold significantly as observed in CD spectra and behaves similar to PduA[WT]. We hypothesize that this behavior of protein variants is specific to amino acid residues present in PduA protein and is not correlated to overall topology of the protein.

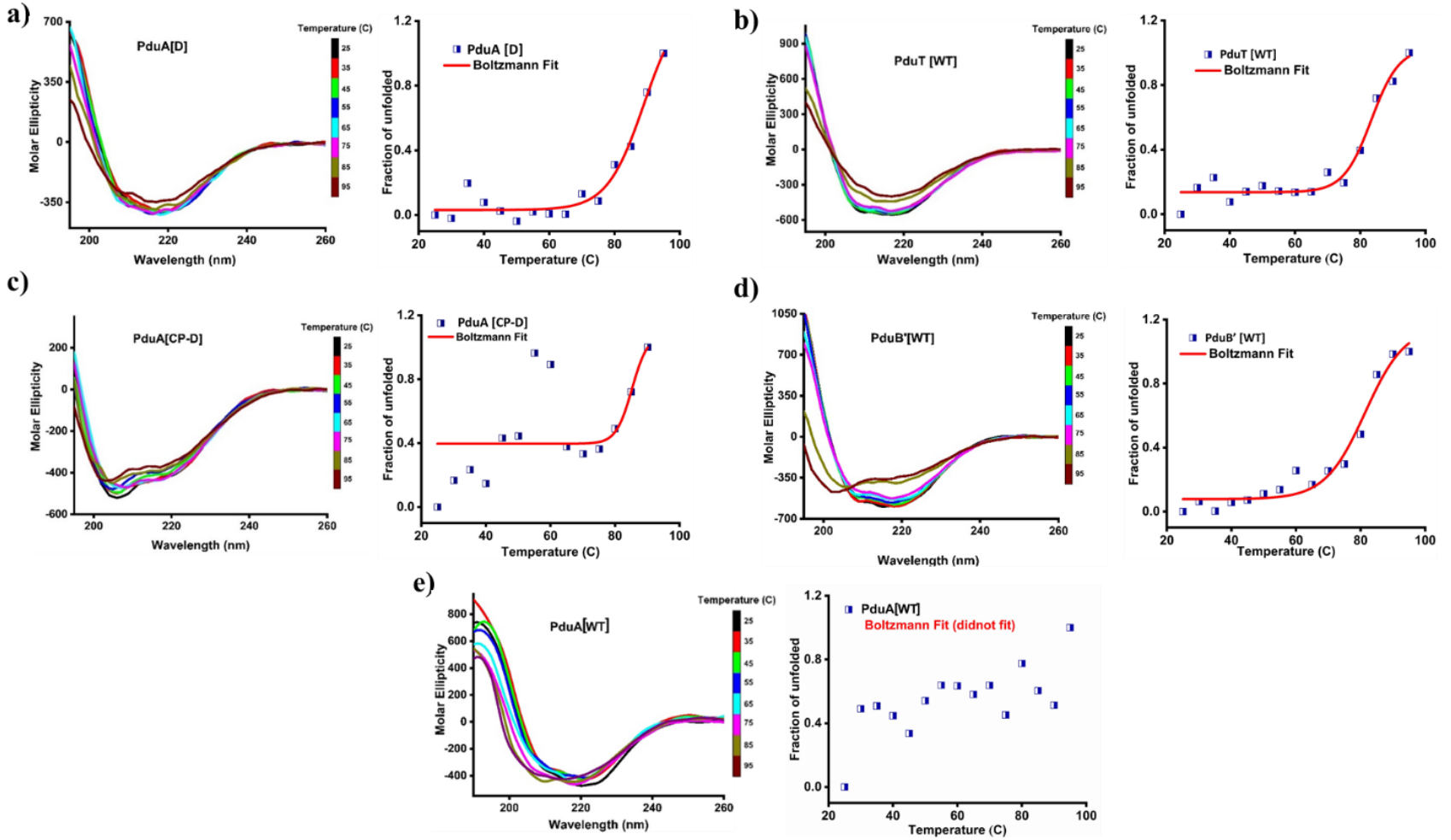


Figure 3.5 Temperature dependent studies of WT and their synthetic variants using Circular dichroism. Circular dichroism spectra showing α-helix secondary structure of the proteins at varying temperatures. a) temperature dependent overlaid spectra of PduA [D] (left) and their fraction of folded plot fit to Boltzmann equation (right). b) temperature dependent overlaid spectra of PduT [WT] (left) and their fraction of folded plot fit to Boltzmann equation (right). c) temperature dependent overlaid spectra of PduA [CP-D] (left) and their fraction of folded plot fit to Boltzmann equation (right). d) temperature dependent overlaid spectra of PduB' [WT] (left) and their fraction of folded plot fit to Boltzmann equation (right). e) temperature dependent overlaid spectra of PduA [WT] (left) and their fraction of folded plot which did not fit to Boltzmann equation (right). Boltzmann fits are in red for visual clarity. The color bar represents the temperature gradient from 25 ^{o}C (black) to 95 ^{o}C (olive green) with an interval of 10 ^{o}C.

3.6 Temperature dependent size distribution studies of the protein fold variants using DLS

Next we proceeded and performed temperature dependent DLS studies to measure the size variation of protein higher order assemblies with respect to temperature ramp. For temperature dependent studies we measured the number mean percentage to avoid the population due to aggregation of large sizes assemblies present in micron size range. The hydrodynamic diameter was recorded while subjecting the proteins to higher temperature till 95 ^{o}C from 25 ^{o}C. We compared all the dimer protein variants. PduA [D] protein variant showed a constant size range around 300nm [Figure 3.6 (b)]. There is no significant decrease in size observed with temperature ramp. This is probable due to the reason that PduA [D] forms tubes as supramolecular assemblies which are held with strong interactions and doesn't break easily into smaller structures. On the other hand, PduT [WT] [Figure 3.6 (a)] which is an equivalent protein to PduA [D] in topology but assembles as thin sheets shows a significant decrease in size from 250nm to 150nm starting from 55 ^{o}C. PduB' [WT] also significantly showed a decrease in size from 600nm to 200nm on heating after 55 ^{o}C. This suggests a disassembly of protein sheets formed by PduB' [WT] [Figure 3.6 (c)] protein into smaller size. Interestingly, PduA [CP-D] protein which is structurally similar to PduB' [WT] showed a decrease in size distribution from 200nm to 150nm [Figure 3.6 (d)] with increase in temperature. When compared to PduA [WT] [Figure 3.6 (e)] protein, it showed decrease in size from 400nm to 200nm on heating from 25 ^{o}C to 90 ^{o}C. There is a disassembly of sheets formed by all protein variants at and above 55 ^{o}C but a significant size shift is observed in PduB'[WT] and PduT[WT] pseudohexamer proteins. PduA [D] which assembles as tubes showed no change in the size. This temperature, 55 ^{o}C, where disassembly is observed for the protein variants is significant when the stability of MCP is taken into picture. The MCP is stable till 55 ^{o}C where it loses its diol dehydratase activity to 50% [112].

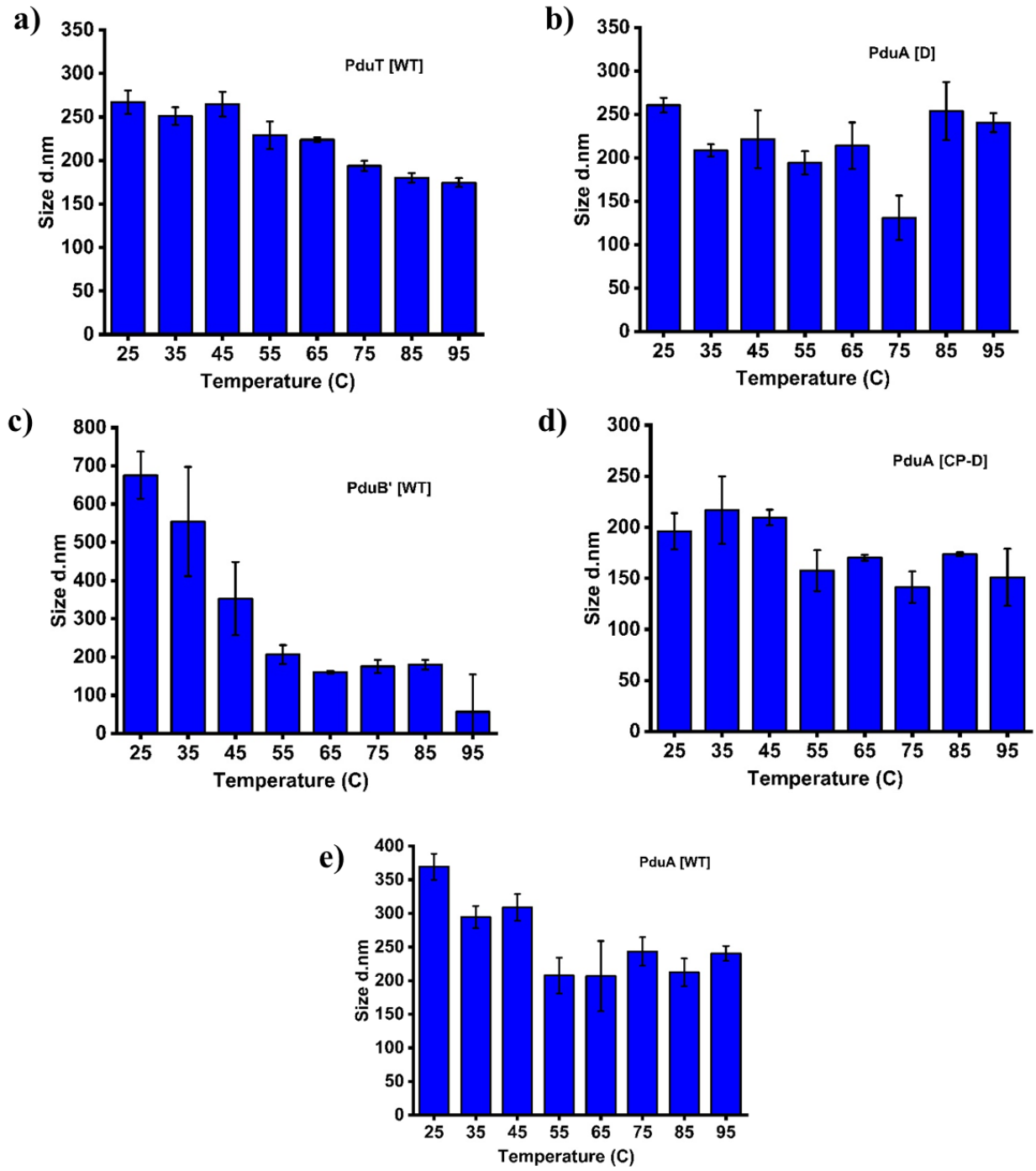


Figure 3.6 Temperature dependent size distribution studies of WT and their synthetic variants using Dynamic Light Scattering (DLS). Bar graphs showing the temperature dependent variation in sizes using Dynamic Light Scattering (DLS). a) Bar graph shows the decrease in the size from 250nm to 150nm with increase in temperature for PduT [WT]. b) Bar graph shows the constant size distribution in the range of 300nm with increase in

temperature for PduA [D]. c) Bar graph shows the exponential decrease in the size distributions from 600nm to 200nm with increase in temperature for PduB' [WT]. d) Bar graph shows the size distribution in the range of 200-150nm with increase in temperature for PduA [CP-D]. e) Bar graph shows the decrease in the size from 400nm to 200nm with increase in temperature for PduA [WT]. Each temperature has three data sets which are represented by overlapping bars with slightly difference in color.

3.7 Fluorescence based temperature dependent stability studies of the protein fold variants using DSF

Further, to compare the global stability of the protein variants we used differential scanning fluorimetry (DSF)[113] to capture the unfolding events at tertiary and quaternary level. This experiment uses an external fluorescent probe, Sypro orange (SO) which binds to an exposed hydrophobic rich patches in a protein and gives fluorescence. Therefore, this will provide information on unfolding of proteins due to thermal assault which will result in alterations in protein higher order structures and unfolding based on the exposed hydrophobic patches which will be probed. Thermal scan of protein domain variants in presence of SO (20x) at 20µM concentrations of protein was carried out from 20 ^{o}C to 100 ^{o}C with 1 ^{o}C increment, equilibrated for 1:30min. The assay was performed in a solution phase. SYPRO orange dye alone was kept as a control in addition to all the protein-dye mixture samples. Obtained DSF curve as Relative Fluorescence Unit (RFU) vs temperature are baseline corrected and transition temperatures are calculated from Boltzmann fit. As a unique signature of shell proteins is already reported that there is higher initial increase in SO fluorescence due to exposed hydrophobic patches on proteins morphology[81]. PduA [WT] protein shows an initial high fluorescence and later enhancement in the fluorescence of SO is observed with an increase in temperature. The transition temperature (T_m) calculated from RFU plot (relative fluorescence unit) for PduA [WT] protein is 59.1 ± 0.1 ^{o}C

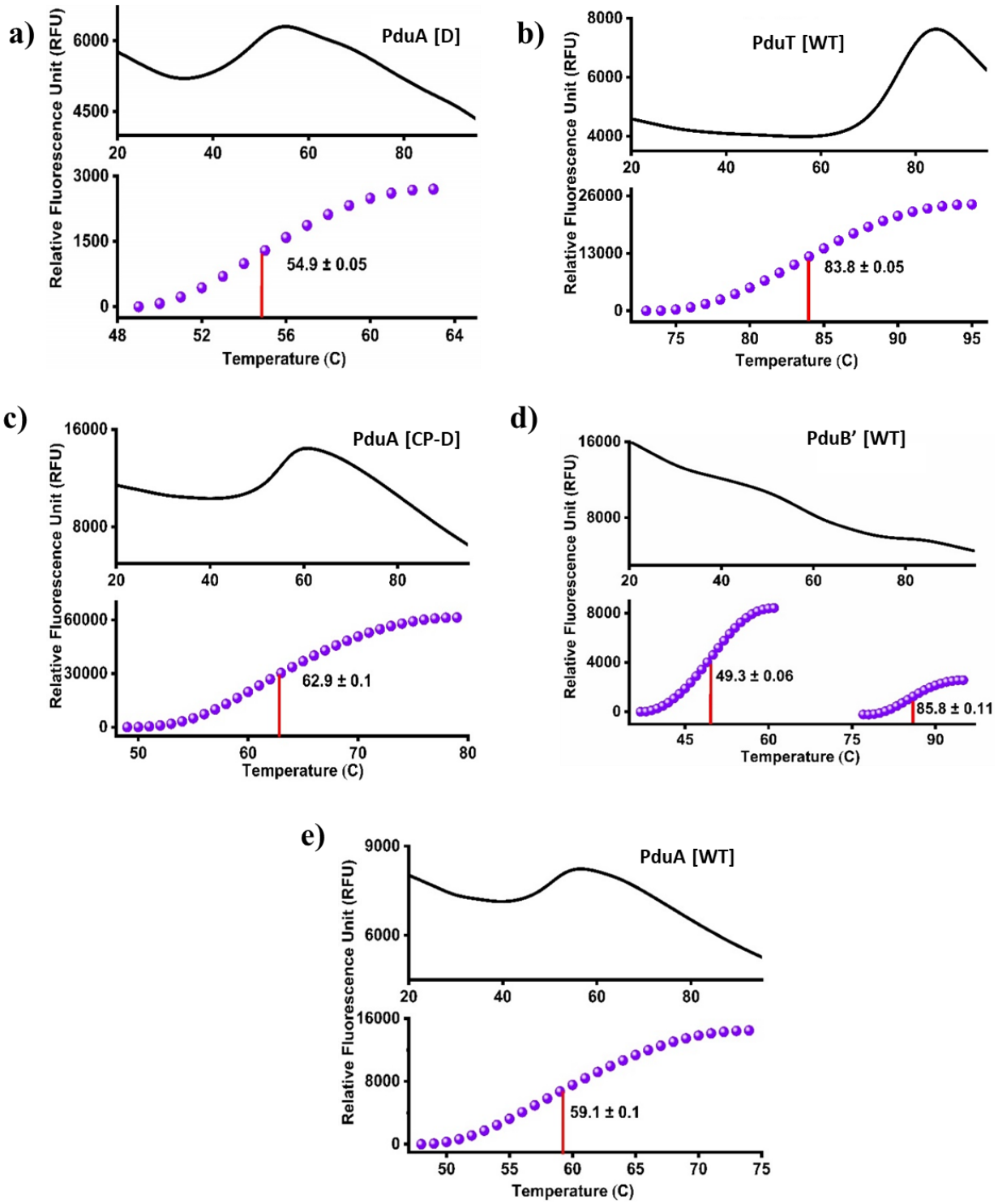


Figure 3.7 Fluorescence based temperature dependent studies of WT and their synthetic variants using Differential Scanning Fluorimetry (DSF). DSF plots monitoring the fluorescence emitted by Sypro Orange on detecting exposure of hydrophobic patches upon

thermal ramping from 20 °C to 90 °C. DSF data is plotted as RFU vs temperature. a) Raw RFU curve (top), Boltzmann fit for transition temperature after baseline correction (bottom) resulting in Tm 54.9 ± 0.05 °C for PduA [D]; b) Raw RFU curve (top), Boltzmann fit for transition temperature after baseline correction (bottom) resulting in Tm 83.8 ± 0.05 °C for PduT [WT]; c) Raw RFU curve (top), Boltzmann fit for transition temperature after baseline correction (bottom) resulting in Tm 62.9 ± 0.1 °C for PduA [CP-D]; d) Raw RFU curve (top), Boltzmann fit for transition temperature after baseline correction (bottom) resulting in two different Tm 49.3 ± 0.06 °C and Tm 85.8 ± 0.11°C for PduB' [WT]; e) Raw RFU curve (top), Boltzmann fit for transition temperature after baseline correction (bottom) resulting in Tm 59.1 ± 0.1 °C for PduA [WT]. Tm values are marked by a red vertical line. DSF curves shown here are representation of three independent experiments.

[Figure 3.7 (e)]. Further, DSF curves were also de-convoluted to multiple peak fitting due to broad peaks observed, indicating multiple transitions. The broad peak when deconvoluted is best fitted to two peaks to 56.2 ±0.2°C and 69.1 ±1.4°C [Figure 3.8 (e)]. The peaks around 50-60°C correlates to the destabilization of BMC shell and reduced BMC activity. We expected an increase in T_m for dimer protein variants of PduA [WT] but interestingly, similar characteristic was observed with PduA [WT] protein variants i.e., PduA [D] at 54.9 ± 0.05 °C [Figure 3.7 (a)] and PduA [CP-D] at 62.9 ± 0.1 °C [Figure 3.7 (c)]. There is a slight increase in T_m in circularly permuted dimer version of PduA [WT]. However, on deconvolution PduA[D] is best fitted to three peaks 53.1 ±2.7°C, 67.7 ±4.1°C, 83.0 ± 3.1°C and PduA[CP-D] fit to two peaks 61.3±0.4°C and 76.6±0.9°C [Figure 3.8 (a,c)]. PduT [WT] pseudohexmer protein has shown high stability with Tm observed at 83.8 ± 0.05 °C and multiple peak fit to single peak fit at 82.9±0.05°C [Figure 3.7 (b), 3.8 (b)]. PduB' [WT] pseudohexamer was also observed with high T_m but with another transition at lower temperature. The two transitions at 49.3 ±0.06 and 85.8 ±0.1 °C [Figure 3.7 (b)] and on deconvoluting peaks, the two transition temperatures were observed at 51.8±1.6°C, 82.3±0.3°C [Figure 3.8 (b)]. This behavior of PduB' [WT] was interesting which was observed earlier as well indicating possibility of different population stabilities[81].

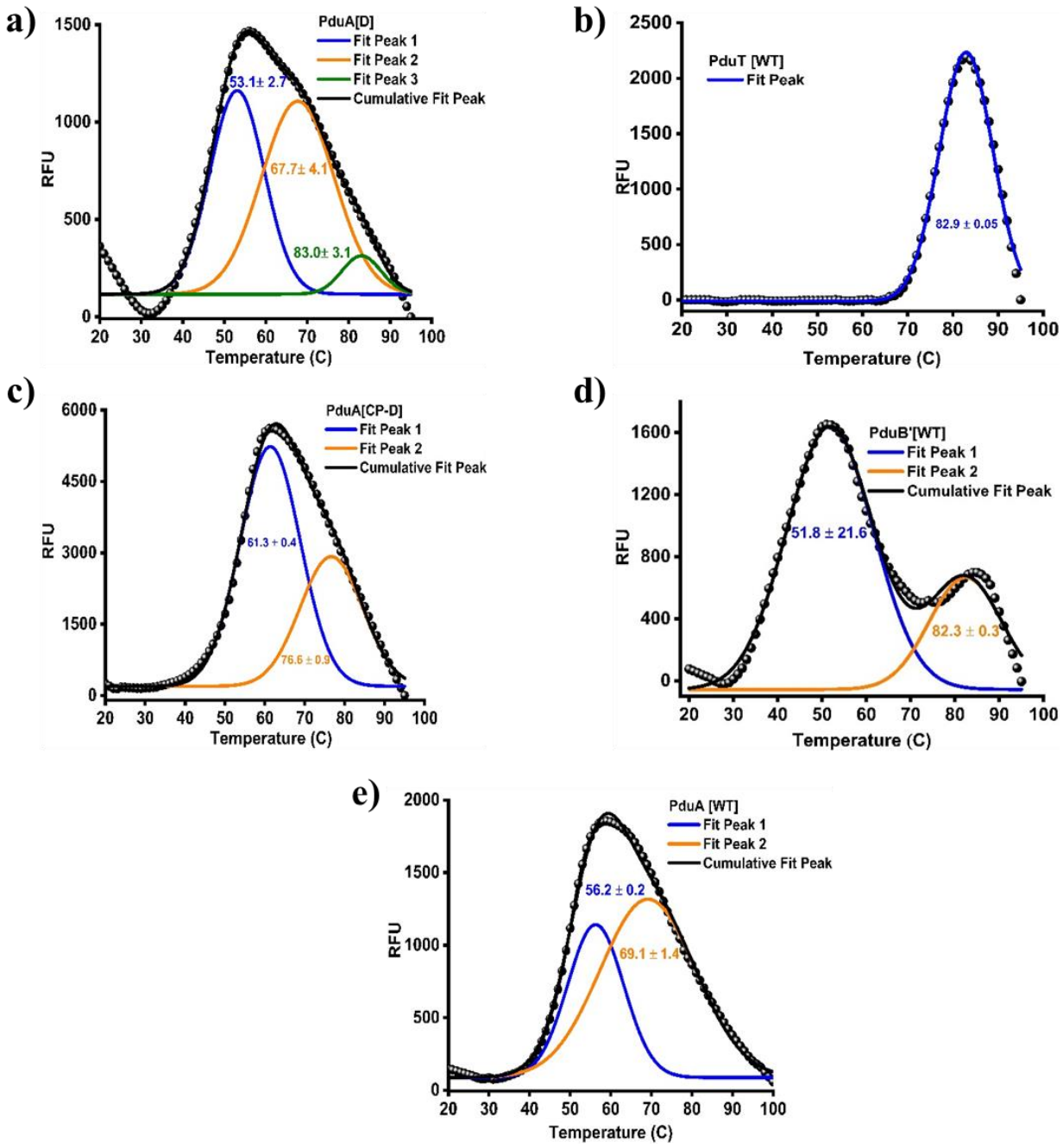


Figure 3.8 Multiple peak fitting of DSF curves for WT and their synthetic protein variants. DSF curve is deconvoluted and fitted to multiple peak transitions. a) Transition temperature after multiple peak fitting resulting in three peak transitions at 53.1 ±2.7^{o}C, 67.7 ±4.1^{o}C, 83.0 ± 3.1^{o}C for PduA [D]; b) Transition temperature after multiple peak fitting resulting in single

peak fit at 82.9±0.05°C for PduT [WT]; c) Transition temperature after multiple peak fitting resulting in two peaks at 61.3±0.4°C and 76.6±0.9°C for PduA [CP-D]; d) Transition temperature after multiple peak fitting resulting in two peaks at 51.8±1.6°C, 82.3±0.3°C for PduB' [WT]; e) Transition temperature after multiple peak fitting resulting in two peaks to 56.2 ±0.2°C and 69.1 ±1.4°C for PduA [WT]. Transition temperatures are mentioned at the peak positions. DSF curves with multiple peak fitting shown here are representation of three independent experiments.

Table 3.2 Transition temperatures for dimer domain protein variants obtained from Differential Scanning Fluorimetery, determined by Boltzmann fit and multiple peaks fitting and Circular Dichroism determined by Boltzmann fit.

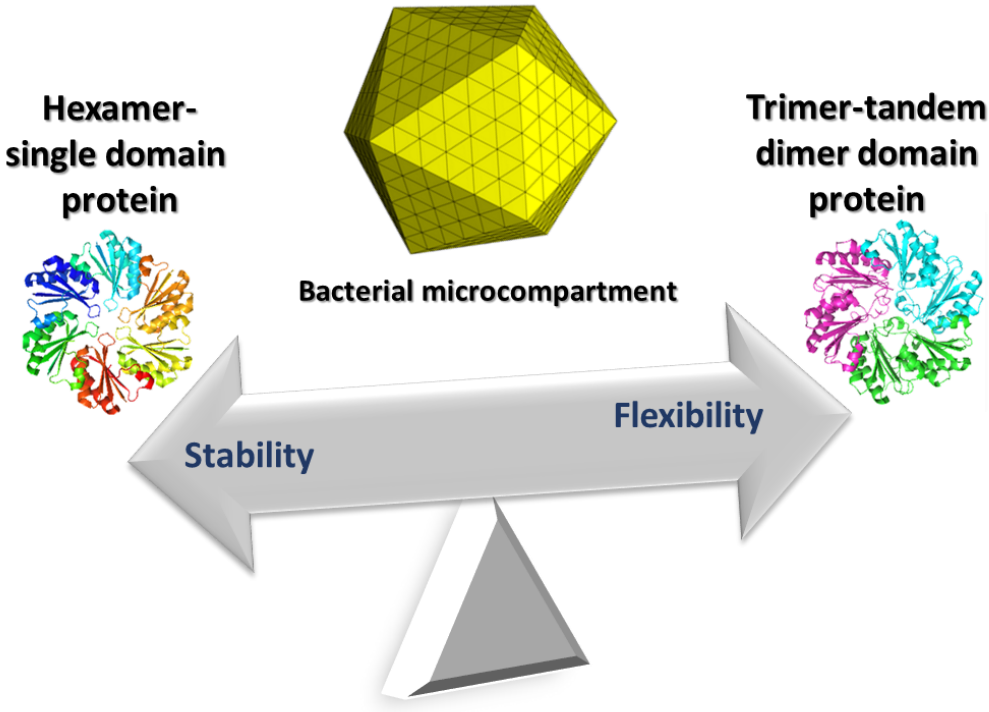
	CD (Boltzmann fit)	DSF (Boltzmann fit)	DSF (Multiple peak fit)
PduB' [WT]	81.4 ± 2.4°C	49.3 ±0.06 °C 85.8 ±0.1 °C	51.8±1.6 °C 82.3±0.3 °C
PduT [WT]	83.4 ± 1.6°C	83.8 ± 0.05 °C	82.9±0.05 °C
PduA[D]	89.01 ± 4.5°C	54.9 ± 0.05 °C	53.1 ±2.7 °C 67.7 ±4.1 °C 83.0 ± 3.1 °C
PduA[CP-D]	85.2 ± 9.3°C	62.9 ± 0.1 °C	61.3±0.4 °C 76.6±0.9 °C
PduA [WT]	-	59.1 ± 0.1°C	56.2 ±0.2 °C 69.1 ±1.4 °C

3.8 Conclusion

This work sheds light on the significance of existence of two type of oligomerization states of proteins present in the microcompartment shell. In an attempt, two synthetic dimer protein variants were generated using a sequence of one of the major shell proteins PduA[WT] which is a hexamer. We compared a dimer variant of PduA[WT] with its native counterpart PduT[WT] which is structurally similar to it, and a circularly permuted dimer variant of PduA[WT] which is structurally equivalent to PduB'[WT]. We found interesting results where foremost, the proteins after purification were folded in their native conformations with presence of an α-helix secondary structure as assayed by circular dichroism and fluorescence experiments. Further, studies by DLS and TEM, evidently proven that they retained the self-assembling property and forms higher order supramolecular structures. All the shell proteins forms sheets whereas PduA [D] forms tubes, probably due to the bent angle formed between the hexamers due to dimerization. Temperature dependent results show that the protein variants are thermostable (Table 3.2). However, native dimer proteins PduB' and PduT are found to be stable till 80 ^{o}C and DLS data showed the disassembly of higher order structures into smaller size at 50 ^{o}C at which MCP loses its 50% activity. Thus, disassembly at this temperature abrogates the assembly of shell proteins of MCP as reported earlier. However, no significant disassembly is observed with protein variants. On comparing PduA [WT] and its synthetic variants using thermal melt studies shows that the secondary structure elements in PduA sequence protein variants did not unfold till 90 ^{o}C and indicates they are compact and rigid as native PduA [WT] protein even though they share a topology similar to native dimer proteins. From these studies we can infer that the protein sequence governs the biophysical properties of BMC domain proteins and consecutively their functions as well. Irrespective of the topology attained the properties belong to the original core of the amino acid sequence. This defines why BMC has evolved with a particular sequence and adopted a specific conformational topology for hexamer and psuedohexamers which defines their physiochemical properties. Sequence plays a more important role in BMC domain proteins which balances and provides both rigidity and flexibility to BMC to function better in a coordination manner. This finding will further aid in studying the functional attributes of the BMC proteins.

CHAPTER – 4

Understanding the role of single-domain proteins by dissecting tandem domains in MCP



4.1 Introduction

Bacterial microcompartment shell proteins are known to oligomerize from single-domain or dimer-domain as hexamer and pseudo hexamers respectively. In the previous chapter, PduA [WT] protein sequence which is a single domain was explored to form the dimer protein variants. In one instance, a dimer protein was generated by directly fusing two monomers in a tandem fashion with the help of a linker present in the native PduB' [WT] sequence. This variant is similar to PduT [WT] protein already present in BMC. In another instance, a circular permutant dimer is generated from PduA [WT] similar to PduB' [WT] in overall topology. The behavior of these dimer variants was found to be sequence-specific, not structure-specific. This result was unexpected and seems interesting which leads to another question what if all shell proteins present in BMC are only single-domain proteins oligomerizing as hexamers? How will the sequence and structure now determine the structure and function of overall BMC? A strong correlation between sequence specificity and topology is witnessed which forced us to look at another picture of the existence and role of single-domain shell proteins in the microcompartment. Single-domain proteins that were experimentally verified to be present in the 1,2 Pdu microcompartment which oligomerizes as hexamers from six monomer units are PduA, PduJ, PduU [41,42,64]. PduJ has an 80% identical amino acid sequence to PduA and it is known to play an essential role in MCP assembly in coordination with PduA protein[64,111]. PduU is a circular permutant version of PduA canonical fold which is a minor component in the MCP shell. The relatedness of shell proteins by sequence and topology has been fascinating. Also, we are aware of the fact that conformation and sequence variation in proteins dictate their function[114,115]. So, we wondered what if the sequence from the dimer protein is used to generate a single-domain protein? The study of their properties will shed light on the interesting facts when compared to native single-domain proteins. So, we focused on generating a few single-domain protein variants from one of the major shell proteins i.e., PduB' [WT]. Therefore, in this chapter, I explored PduB' [WT] dimer-domain protein to get single-domain proteins and study their properties and compare them to native single–domain protein PduA [WT]. PduB' [WT] was dissected into two individual single-domain proteins i.e., PduB' [D1] and PduB' [D2]. Studying these single-domain proteins has resulted in some interesting results owing to their independent existence. Further, we thought to get a single domain that has a

Protein variants	Amino acid sequence
PduB' [WT]	MAEKSCSLTEFVGTAIGDTLGLVIANVDTALLDAMKLEKRYRSIGILGARTGAGPHIMAADEAVKATNTEVVSI ELPRDTKGGAGHGSLIILGGNDVSDVKRGIEVALKELDRTFGDVYGNEAGHIELQYTARASYALEKAFGAPIGR ACGIIVGAPASVGVLMADTALKSANVEVVAYSSPAHGTSFSNEAILVISGDSGAVRQAVTSAREIGKTVLATLG SEPKNDRPSYI
PduB' [D1]	MAEKSCSLTEFVGTAIGDTLGLVIANVDTALLDAMKLEKRYRSIGILGARTGAGPHIMAADEAVKATNTEVVSI ELPRDTKGGAGHGSLIILGGNDVSDVKRGIEVALKELDRTFG
PduB' [D2]	GHIELQYTARASYALEKAFGAPIGRACGIIVGAPASVGVLMADTALKSANVEVVAYSSPAHGTSFSNEAILVIS GDSGAVRQAVTSAREIGKTVLATLGS
PduB' [CP-D2]	RACGIIVGAPASVGVLMADTALKSANVEVVAYSSPAHGTSFSNEAILVISGDSGAVRQAVTSAREIGKTVLATL GSGHIELQYTARASYALEKAFGAPIG
PduA [WT]	MQQEALGMVETKGLTAAIEAADAMVKSANVMLVGYEKIGSGLVTVIVRGDVGAVKAATDAGAAAARNVG EVKAVHVIPRPHTDVEKILPKGISQ

Figure 4.1 Amino acid sequences of the desired single-domain protein constructs for the present study. PduB' [WT] (top most), PduB' [D1] (second from the top), PduB' [D2] (middle), PduB' [CP-D2] (second from the bottom), and PduA [WT] (bottom).

similar 3D topology like PduA [WT] and sequence from PduB' [WT]. To obtain this construct we used PyMOL software[87] and performed structural alignment of monomer units from PduA[WT] to PduB' [WT]. PduA[WT] aligns with the second domain of PduB' with a better RMSD score of 0.85 Å than the first domain with a RMSD of 1.49 Å [Figure 4.3 (c)]. So, we, chose the second domain of PduB'[WT] and circularly permuted it to get a structure similar to native PduA and characterized. It is referred to as PduB' [CP-D2] in future text. The sequence of proteins used is shown in Figure 4.1.

Here we are using PduB'[WT] which is a smaller version of version of PduBB'[WT] protein with an extra N-terminal 37 amino acid residues[61]. These 37 amino acids form intrinsically disordered N-terminus region whose role is established in protein-protein interactions[72,73]. So, in our studies, we explored a smaller version of PduBB' i.e., PduB' and compared both domains of PduB' protein (domain1 and domain2) for its dynamic and flexible behavior. As mentioned in Chapter 3, PduB'[WT] is noticed to carry a less compact structure than PduA[WT] as observed from temperature-dependent circular dichroism studies. We speculate this is due to the presence of two domains which governs the overall less compact structure and melts at high temperature. To understand the probable behavior from each domain we first used an online web server, MEDUSA[92] which predicts protein flexibility from amino acid sequence. We fed amino acid sequences from single domains of PduB'[WT], PduB' [D1], and PduB' [D2] [Figure 4.2(a)] in the online server. Along with these slice-up domains, other shell proteins were also analyzed. PduA [WT] single-domain hexamer protein and PduN [WT] single-domain pentamer protein were compared. The results showed that PduB' [D1] and PduN [WT], proteins have high flexibility percentages of 43 % and 41% respectively. This data was very interesting knowing that the PduB' [D1] has an equal percentage of flexibility as PduN [WT] protein which is a pentamer protein and known to be the most flexible protein[55,74]. However, PduB' [D2] has 28% flexibility comparably less than PduB' [D1]; and is similar to the flexibility percentage of PduA [WT] protein (31%). Interestingly, PduK [WT] protein also has a comparable flexible score of 43% to PduB' [D1]and PduN [WT]. PduK protein structure and function are not known but it comprises a single BMC domain with a long extended C-terminal which is a disordered region that might be responsible for high flexibility[42,71]. PduT [WT] protein which is another pseudohexamer like PduB', was also analyzed as two domains separately. PduT [WT] domain 1 and domain 2 have 26% and 22%

flexibility respectively. Other monomer shell proteins PduU [WT] and PduJ [WT] have 33% and 29% flexibility (Table 4.1). We also verified these results using another online server, CABSflex[91] which predicts the protein flexibility from protein structure. It generates protein dynamics using a coarse-grained model. We predicted the protein structures for this simulation using alphafold2[89,90], and visualized and analyzed them using PyMOL[87].

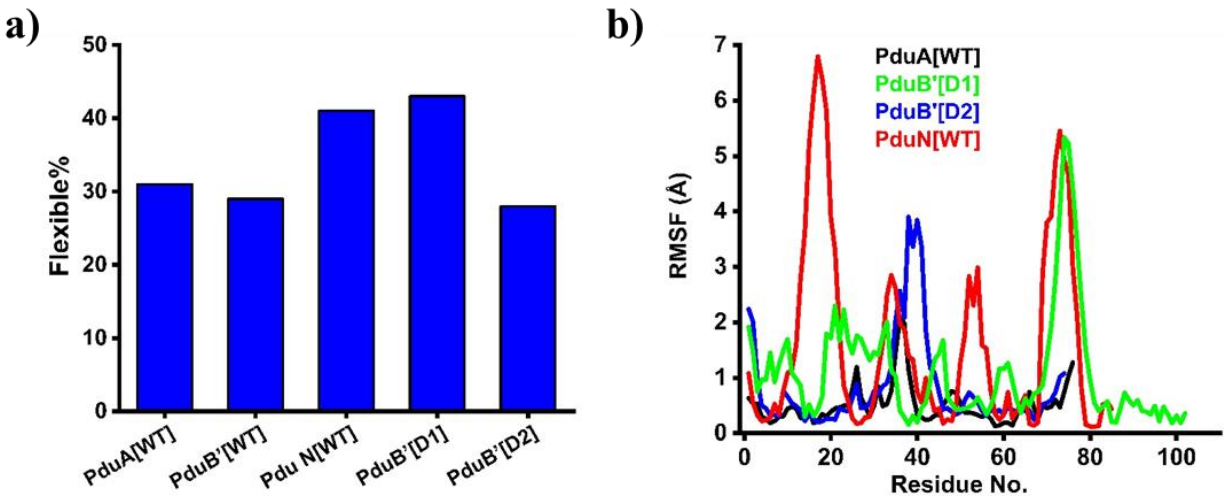


Figure 4.2 a) Protein flexibility prediction using MEDUSA web-server for PduA[WT], PduB'[WT], PduN[WT], PduB'[D1] and PduB'[D2] b) Residual flexibility plot using CABS-Flex web-server. CABS-Flex estimated RMSF scores for the flexibility and rigidity of PduA[WT], PduN[WT], PduB'[D1] and PduB'[D2].

On comparing proteins, we found a remarkable flexibility in PduB' [D1] and PduN [WT] shell proteins having high mean residual fluctuation (rmsf) values at many stretches along the stretch of protein sequence [Figure 4.2(b)]. Encouraging by these results, we wanted to study these proteins further using an array of experimental biophysical techniques. We first initiated by generating the constructs of PduB'[D1] and PduB' [D2] as individual single-domain proteins, and a circularly permuted PduB' domain2 named PduB' [CP-D2] using DNA recombinant approaches. Here, PduB' [CP-D2] will serve as a structurally equivalent protein of PduA [WT] which will share the

same 3D topology. Studying domain 1 and domain 2 of PduB' will answer whether the flexibility in a structure as seen by circular dichroism is a cooperative effect of both domains or is a property of a single domain present. Also, the existence of a single domain of all three proteins and studying their properties will be interesting. This will also put some insights into the question that why a particular sequence exists as a dimer protein or as a single-domain protein.

Table 4.1 Table showing flexibility percentage values when compared domain flexibility using MEDUSA server.

	Rigid%	Flexible%
PduB' [D1]	57	43
PduB' [D2]	72	28
Pdu N [WT]	59	41
PduT [D1]	74	26
PduT [D2]	78	22
Pdu J [WT]	71	29
Pdu K [WT]	58	43
Pdu U [WT]	67	33
Pdu A [WT]	69	31

We also performed the multiple sequence alignment of all the single-domain proteins to get insights into the relatedness of the protein sequences. PduB' [D1] and PduB' [D2] show 19% similarity (Figure 4.3 (e)). On comparing the sequence similarity of PduA [WT] with both the domains, PduB'[D1] and PduB' [D2], it is seen to have 16% and 17% sequence similarity. Rearranged protein sequence on circular permutation, PduB' [CP-D2], the similarity score shifts to 19%. This suggests that domain 2 is more related and has more sequence similarity to PduA

[WT] and on rearrangement, it increased further. When compared to the structures, PduB' [D1] and PduB' [D2] structures are aligned at RMSD of 5.32Å. PduB' [D1] and PduA [WT] aligned at RMSD 1.49Å. PduB' [D2] and PduA [WT] aligned at RMSD 0.85Å. After circular permutation, PduB' [CP-D2] and PduA [WT] aligned at RMSD 0.93Å. [Figure 4.3 (a-d)].

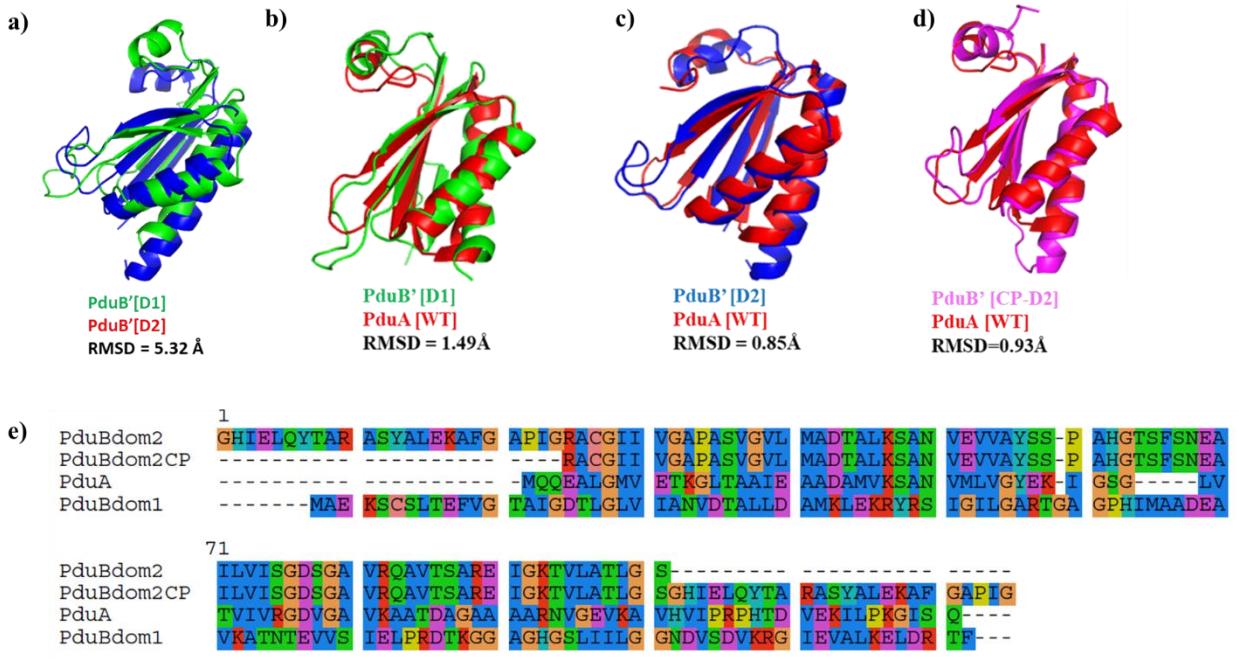


Figure 4.3 Structural and sequence alignment of single-domain protein variants. a) PduB' [D1] and PduB' [D2] aligned at RMSD 5.32Å. b) PduB' [D1] and PduA [WT] aligned at RMSD 1.49Å. c) PduB' [D2] and PduA [WT] aligned at RMSD 0.85Å. d) PduB' [CP-D2] and PduA [WT] aligned at RMSD 0.93Å. Structure alignments were performed using PyMOL. e) Multiple sequence alignment of single–domain protein variants using *ClustalW*.

4.2 Generation of Single-domain protein variants

PduB' WT protein was dissected into its domains from the linker region joining the two domains. The first domain was dissected to the first half till the linker region which corresponds to 350 residues, PduB' [D1], and domain 2 was dissected to the second half after the linker region and contributes to 300 residues, PduB' [D2]. These domain residues were then used as inputs to predict the structure of proteins from alphafold2 [Figure 4.4 (a-i, iii)]. Further, to generate the circular permutant protein of PduB' [D2], we first aligned it with PduA monomer using the PyMOL software. The aligned structures were then used to find the circular permutation. [Figure 4.4 (a-ii)] shows the region marked by a scissor which will be cut to create new N and C- terminals and a region that will be joined, highlighted by an orange curved line. After having a circular permutant variant, the protein structure was predicted by alphafold2 and visualized by using PyMOL software [Figure 4.4 (a-iv)]. To generate these single-domain variants, we used the Gibson assembly protocol as a DNA recombinant method. The Gibson assembly method is described explicitly in Chapter 3 as the same method was used to generate dimer protein variants. The design of the circularly permuted variant from PduB' [D2] to generate a construct is shown in [Figure 4.4 (b)]. The amplified DNA fragments corresponding to each protein sequence were ligated to a linearized vector using the Gibson assembly[88] method and transformed in the DH5α cloning strain. The positive clones were screened using the colony PCR methodology [Figure 4.4 (c)]. The positive clones were then verified with Sanger sequencing and transformed into expression strain BL21(DE3) PlysS cells. The protein expression at 0.5mM IPTG concentration for 8 hr at 28 ^{O}C was standardized as described in Chapter 3 and protein was obtained in a soluble fraction. Ni-NTA affinity chromatography was used for protein purification as previously used for WT shell proteins. The lysate of the protein mixture obtained after cell lysis was passed through the pre-equilibrated Ni-NTA resin, equilibrated with column buffer (50 mM Tris, 200 mM NaCl, and 5 mM Imidazole, pH 7.5). The column was then washed with the same buffer and then increasing the concentration of Imidazole to 50 mM. The purified proteins were eluted at 200 mM Imidazole concentration and protein concentration was assayed using Bradford assay methodology.

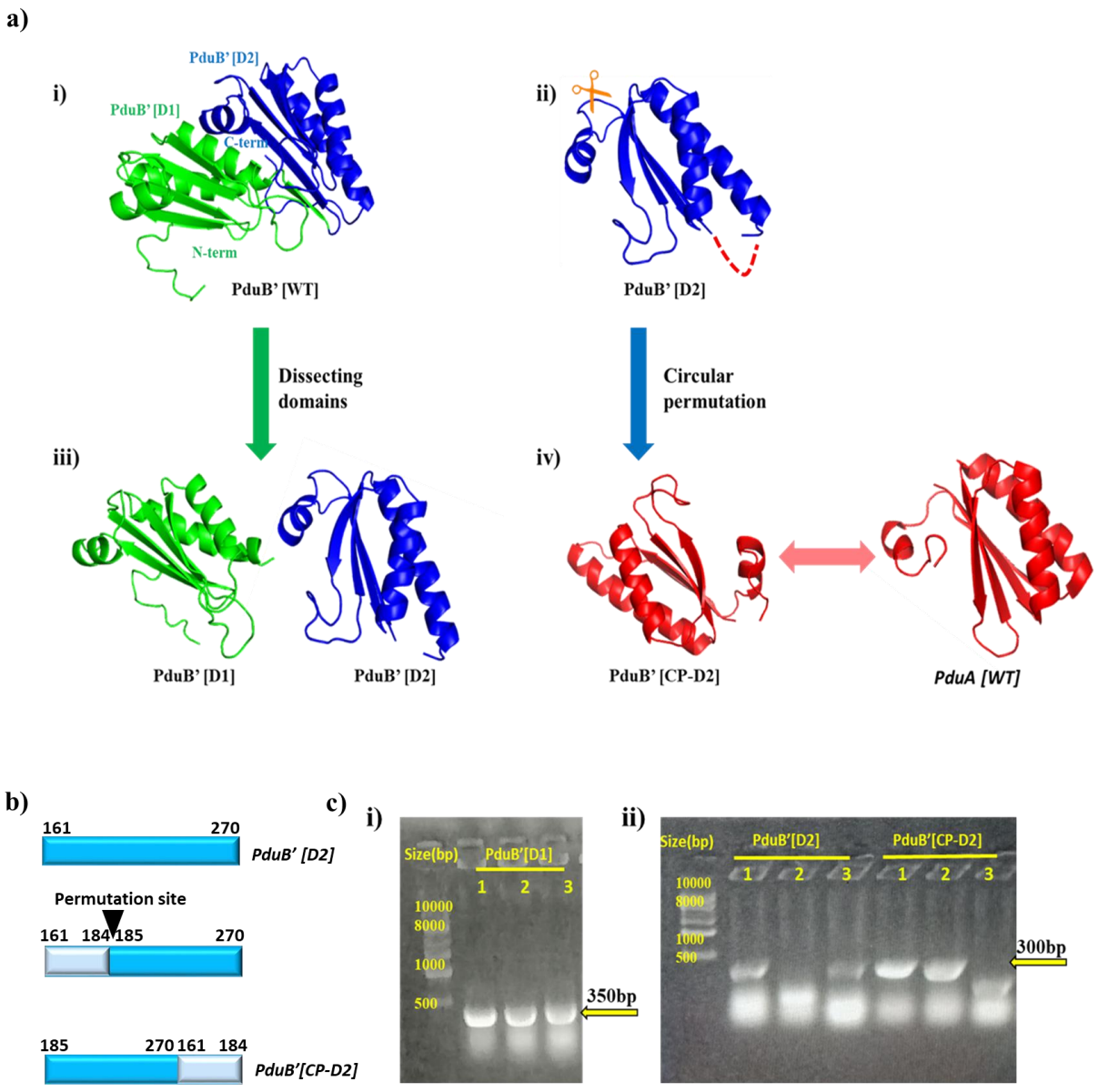


Figure 4.4 Schematics for generation of single-domain protein variants. a) (i) PduB'[WT] modelled from PDB 4FAY. Each domain is shown in a different color. PduB' [D1] (green) and PduB' [D2] (blue). (ii) PduB' [D2] showing the two ends which will be joined by a linker and a scissor marking the region where new N and C terminus will be generated (blue). (iii) Dissected individual domain structures PduB' [D1] and PduB' [D2] are predicted from AlphaFold2 and visualized using PyMOL (iv) Image showing the circularly permuted PduB' [D2] (red) and PduA[WT]. Image of PduB' [CP-D2] generated using PyMOL which was predicted from AlphaFold2. (b) Schematic presentation of the design of single-domain

protein variants from PduB'[WT]. For circularly permuted variants, the permuted site is highlighted with a black arrow along with residue numbers (c) Agarose gel electrophoresis image of cloning of single-domain variant proteins. (i) Colony PCR products screening for positive clones of PduB' [D1]. (ii) Colony PCR products screening for positive clones of PduB'[D2] and PduB' [CP-D2].

4.3 Biophysical characterization of natural and synthetic single-domain variants

4.3.1 Circular Dichroism and steady-state fluorescence spectroscopy

To characterize the single-domain variants, we first studied the secondary structure and folding of the protein variants using Far-UV circular dichroism and steady-state intrinsic fluorescence. Far-UV circular dichroism data was recorded for PduB' [D1], PduB' [D2], and PduB' [CP-D2] and compared to PduA [WT] single-domain protein. The Far-UV CD spectra show that all the single-domain protein variants are folded properly into the secondary structure elements [Figure 4.5 (a)] PduA [WT] has characteristic spectra of the α-helix with a dip in ellipticity at 208nm and 222nm. Other single-domain proteins obtained after dissecting PduB' [WT] i.e., PduB' [D1], PduB' [D2], and a circularly permuted single-domain from PduB' [WT] i.e., PduB' [CP-D2] were also found to be correctly folded and gave α-helix spectrum with negative ellipticity at 206nm and 222nm. Further, we used steady-state fluorescence spectroscopy to measure the fluorescence emission of the folded protein elements [Figure 4.5 (b)]. The intrinsic fluorophores (aromatic amino acids) present in the wild-type and the single-domain protein variants are given in Table 4.2. Interestingly, the single domain proteins variants of PduB' [WT] gave fluorescence emission peaks at 327nm, 325nm, and 328nm for PduB' [D1], PduB' [D2], PduB' [CP-D2], and 312 nm for PduA [WT] respectively, when excited at 285nm wavelength. Here, we found a difference in emission peak wavelength in PduA [WT] as compared to other synthetic single-domain variants. This result

suggests different packing of protein structure leading to different emission wavelengths despite sharing similar topology in PduB' [CP-D2] and PduA [WT].

Table 4.2: Table with information about the abundance of aromatic amino acid residues present in shell proteins and their synthetic single-domain variants used in this chapter.

	Name of aromatic amino acids present	Number of aromatic amino acids present
PduA[WT]	Tyrosine Tryptophan Phenylalanine	1 0 0
PduB'[D1]	Tyrosine Tryptophan Phenylalanine	1 0 2
PduB'[D2]	Tyrosine Tryptophan Phenylalanine	3 0 2
PduB'[CP-D2]	Tyrosine Tryptophan Phenylalanine	3 0 2

4.3.2 Dynamic light scattering and TEM analysis to demonstrate the ability of self-assembling of synthetic single- domain protein variants into higher-order structures

Shell proteins own the remarkable property of self-assembling. Also in Chapter 3, we found that the dimer-domain protein variants retained this property. So, we were excited to study the same for single-domain proteins as they were generated by dissecting the dimer domains. On studying

for the ability to self-assemble and form higher-order structures, single–domain proteins also gave interesting results owing to their independent existence. PduB'[D2] showed broad size distribution in the range of 400nm-900nm and TEM images showed long thin sheets [Figure 4.5 (g, h)] whereas PduB'[D1] didn't show any higher order assembled structures [Figure 4.5 (e, f)]. PduB'[CP-D2] showed smaller sheets in TEM images as compared to PduB'[D2] without circular permutation and size ranges from 200-600 nm [Figure 4.5 (i, j)]. It is interesting to note that the second domain of PduB' [WT] with and without circular permutation has resulted in different levels of higher-order organization as observed through TEM analysis. PduB'[CP-D2] assembles into smaller sheets as compared to PduB'[D2]. PduA[WT] has a wide size distribution and thin sheets were observed in TEM images [Figure 4.5 (c, d)].

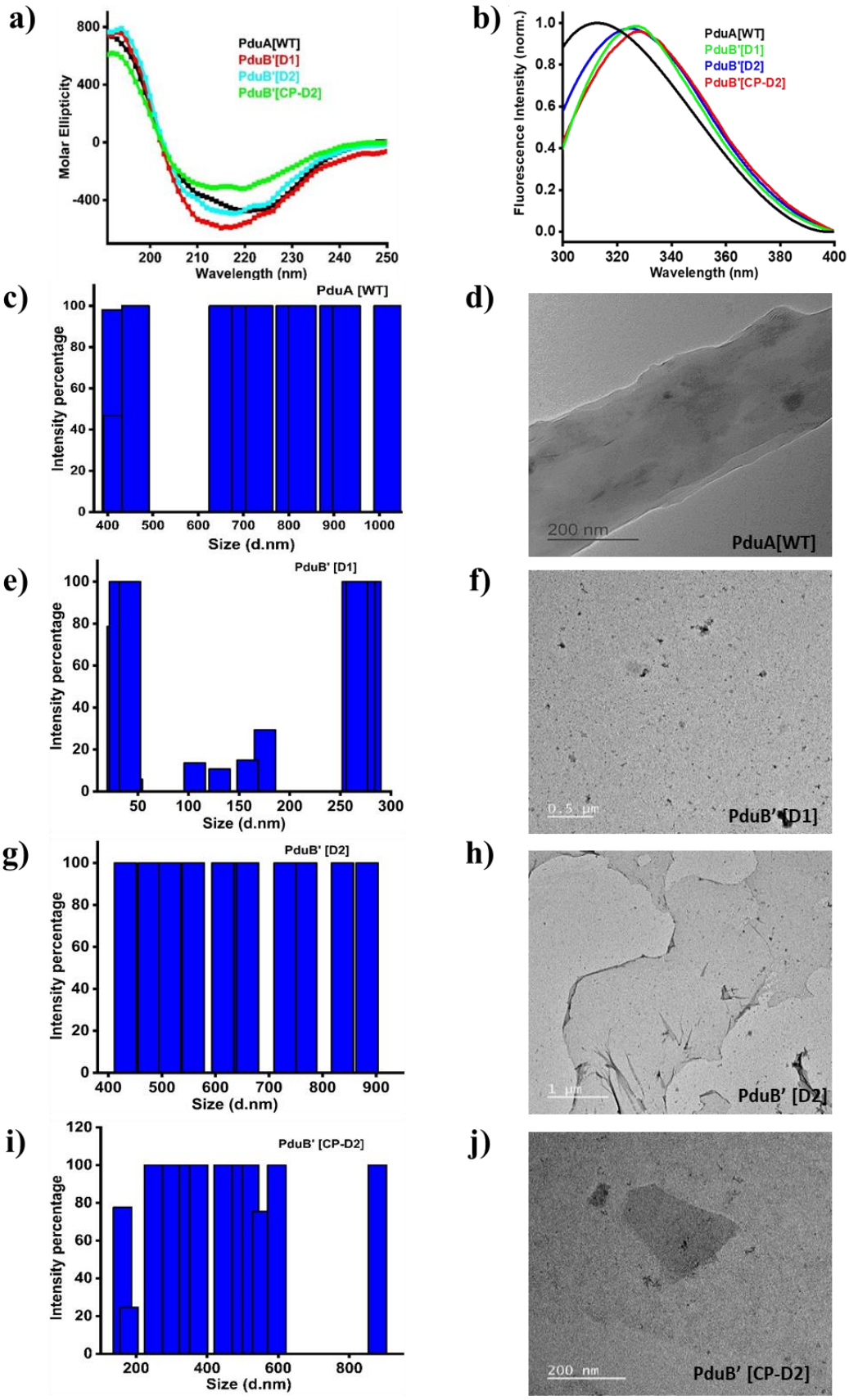

a)
Molar Ellipticity
Wavelength (nm)
PduA[WT]
PduB'[D1]
PduB'[D2]
PduB'[CP-D2]
b)
Fluorescence Intensity (norm.)
Wavelength (nm)
PduA[WT]
PduB'[D1]
PduB'[D2]
PduB'[CP-D2]
c)
PduA [WT]
Intensity percentage
Size (d.nm)
d)
200 nm
PduA[WT]
e)
PduB' [D1]
Intensity percentage
Size (d.nm)
f)
0.5 µm
PduB' [D1]
g)
PduB' [D2]
Intensity percentage
Size (d.nm)
h)
1 µm
PduB' [D2]
i)
PduB' [CP-D2]
Intensity percentage
Size (d.nm)
j)
200 nm
PduB' [CP-D2]

Figure 4.5 Characterization and size-distribution of natural and synthetic single domain protein variants a) Overlapped circular dichroism spectra of PduA [WT] (black), PduB' [D1] (red), PduB' [D2] (blue) PduB' [CP-D2] (green). b) Fluorescence emission spectra overlaid of PduA WT(black), PduB' [D1] (red), PduB' [D2] (blue) PduB' [CP-D2] (green). Dynamic light scattering (DLS) showing size distribution of protein variants in d.nm c) PduA [WT], e) PduB' [D1], g) PduB' [D2], i) PduB' [CP-D2] and TEM micrographs showing self-assembled structures d) PduA [WT], f) PduB' [D1], h) PduB' [D2], j) PduB' [CP-D2].

To investigate further the diverse properties of single-domain protein variants, we performed temperature-dependent studies for all the synthetic single-domain proteins and compared them to natural single-domain variants i.e. PduA [WT]. The temperature-dependent stability studies were performed using circular dichroism to mark the change in the secondary structure of proteins on thermal assault. Further, dynamic light scattering studies were performed to observe the stability and change in the assembly or disassembly of proteins' higher-order structure at high temperatures. In the end, we went to perform fluorescence-based differential scanning fluorimetry studies where we expected to observe the protein's stability at the global level while probing the hydrophobic patches.

4.4 Temperature-dependent stability studies using circular dichroism

We performed temperature-dependent circular dichroism studies with all the single-domain protein variants that we generated. This will provide us with information about the unfolding and uncoiling of secondary structure elements of the proteins with an increase in temperature. We subjected the proteins to a temperature range from 25 ^{o}C to 95 ^{o}C. On analyzing the thermal melt spectra for PduB'[D1], we found no significant change in ellipticity with an increase in temperature [Figure 4.6(a)]. We plotted the fraction of unfolded protein with respect to temperature at 222 nm and fit to the Boltzmann equation (Eq 1).

$$T = \frac{A_1 - A_2}{1 + e^{(T - Tm)/dT}} \quad \text{- Equation (1)}$$

(A_1 and A_2 are the initial and the final values, T_m is the midpoint of the transition, and dT is the slope of the transition)

But we found the absence of a proper S-shaped curve or sigmoid curve [Figure 4.6(b)]. We further found that the second domain with and without CP also doesn't show unfolding when ramped to 90^{o} C as the spectra are overlapping at each temperature and showed the absence of a proper sigmoid curve [Figure 4.6 (c-f)]. Here, dissecting the PduB'[WT] domains has resulted in surprising stability profiles where single domains proteins do not melt easily and form more compact structures with strong interactions similar to PduA[WT] [Figure 4.6 (g, h)].

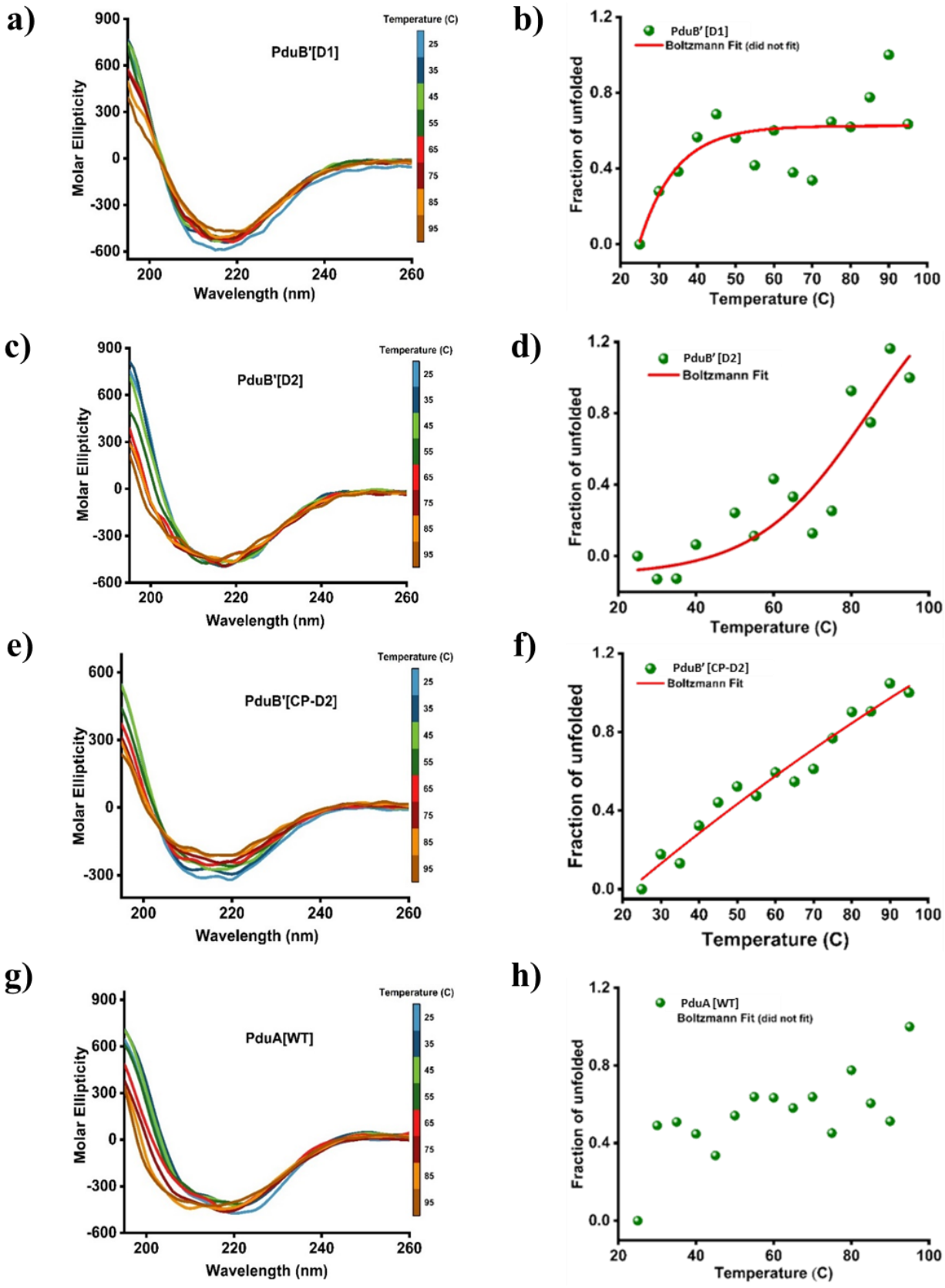

a)
PduB'[D1]
Molar Ellipticity
Wavelength (nm)
Temperature (C)
b)
PduB' [D1]
Boltzmann Fit (did not fit)
Fraction of unfolded
Temperature (C)
c)
PduB'[D2]
Molar Ellipticity
Wavelength (nm)
Temperature (C)
d)
PduB' [D2]
Boltzmann Fit
Fraction of unfolded
Temperature (C)
e)
PduB'[CP-D2]
Molar Ellipticity
Wavelength (nm)
Temperature (C)
f)
PduB' [CP-D2]
Boltzmann Fit
Fraction of unfolded
Temperature (C)
g)
PduA[WT]
Molar Ellipticity
Wavelength (nm)
Temperature (C)
h)
PduA [WT]
Boltzmann Fit (did not fit)
Fraction of unfolded
Temperature (C)

Figure 4.6 Temperature-dependent studies of single domain protein variants using Circular dichroism. Circular dichroism spectra of proteins at varying temperatures. Temperature dependent overlaid spectra of a) PduB' [D1] and their b) fraction of folded plot fit to Boltzmann equation c) PduB' [D2] and their d) fraction of folded plot fit to Boltzmann equation e) PduB' [CP-D2] and their fraction f) of folded plot fit to Boltzmann equation g) PduA [WT] and their h) fraction of folded plot which did not fit to Boltzmann equation Boltzmann fits are in red for visual clarity. The color bar represents the temperature gradient from 25 ^{o}C (blue) to 95 ^{o}C (brown) with an interval of 10 ^{o}C.

4.5 Assessing the thermal stability using dynamic light scattering

Here, similarly to what we performed for dimer-domain proteins, we employed the temperature-dependent DLS experiments for single protein variants with an approach to monitor temperature-dependent changes in the size distribution of higher-order structures. Dynamic light scattering is a great method to estimate the change in the size of a protein with higher-order structures. Any assembly or disassembly of proteins high order structures can be measured. As these protein variants have also the ability to self-assemble and hence are dynamic, we used the number mean as a measure of temperature change. The proteins were subjected to a temperature ramp from 25 ^{o}C to 95 ^{o}C and data was recorded at every 10 ^{o}C difference. On exposing these proteins to higher temperatures we observed there is no significant change in the size of the proteins for PduB'[D2] with and without CP. However, PduB'[D1] protein size shifted from 250nm to 100nm at above 80^{o}C [Figure 4.7(a)]. PduB'[D2] protein does not show any change in the size [Figure 4.7(b)]. PduB'[CP-D2] protein is also stable at higher temperatures so it doesn't show any significant change in size [Figure 4.7(c)]. PduA protein size slightly shifted from 300 to 200 nm at 50C which then remained stable till 95C [Figure 4.7(d)]. This suggests that the higher-order self-assembled structures made by these single-domain proteins are stable towards high temperatures and do not disassemble to a significant extent. However, PduB'[D1] protein transition at 80^{o}C can be related to the melting temperature of PduB'[WT] full-length protein. It signifies that the protein assemblies break into smaller sizes at high temperatures in PduB'[WT] full length containing both the domains and this characteristic is central to domain 1 specifically. No shift in size was observed

in domain 2 of PduB', which implies that it has more stable interactions than that are present in domain 1.

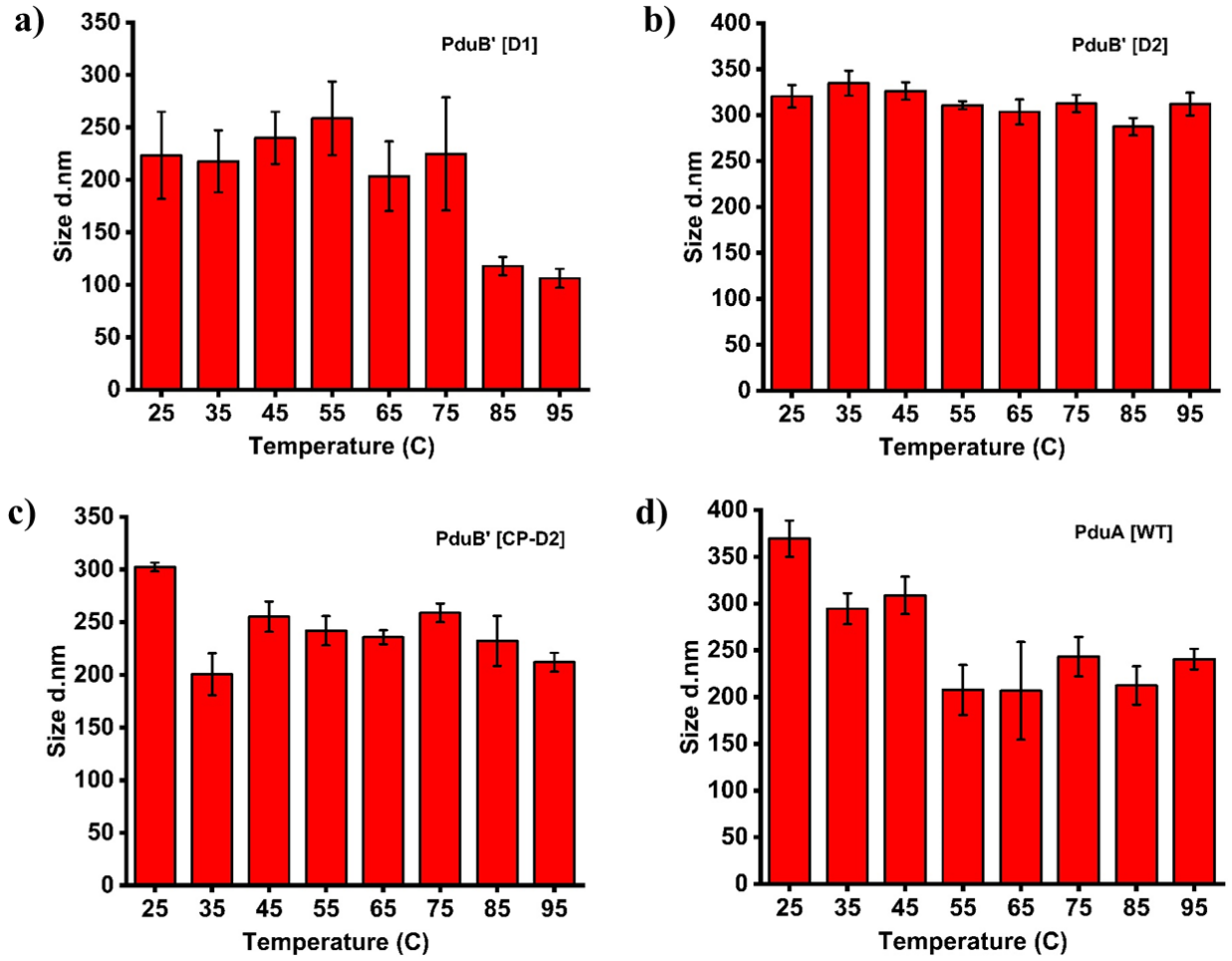


Figure 4.7 Temperature-dependent size distribution studies of single domain synthetic variants using Dynamic Light Scattering (DLS). Bar graphs showing the temperature-dependent variation in sizes using Dynamic Light Scattering (DLS). a) Decrease in the size distributions from 250nm to 100nm with an increase in temperature for PduB'[D1]. b) Constant size distribution in the range of 300nm with an increase in temperature for PduB'[D2]. c) Size distribution in the range of 300-2000 nm with an increase in temperature for PduA[CP-D]. d) Decrease in the size from 400nm to 250nm with an increase in temperature for PduA [WT].

4.6 Thermal Stability measured using fluorescence-based differential scanning fluorimetry

To further assess the temperature-related changes in the protein variants, we used differential scanning fluorimetry to probe the stability of single-domain variant proteins. As already discussed, and mentioned in the principle of DSF technique in chapter 3, briefly here, SYPRO orange dye was used at 20x concentration with 20µM protein concentrations. The temperature was ramped from 20C to 95C with 1 C increment, equilibrated for 1:30min. The increase in SYPRO orange dye fluorescence was probed with the unfolding and increasing exposure of hydrophobic patches in proteins. The RFU vs temperature plots were fit to Boltzmann fit after baseline subtraction to determine the melting temperature (Tm). We have seen in Chapter 3 that PduB' [WT] protein displays two transition temperatures in the DSF experiment at 49.3 ± 0.06 and 85.8 ±0.1 and on deconvoluting peaks 51.8±1.6^{o}C, 82.3±0.3^{o}C. On dissecting the domains in this chapter, we found the Tm of each domain, folded and self-assembled into tertiary and quaternary architect protein. T_m for PduB'[D1] as calculated from Boltzmann fit is 57.6 C ± 0.1^{o}C [Figure 4.8 (a)]. Interestingly, it is important to note here is the RFU plot of PduB'[D1] which is similar to PduB'[WT] [inset Figure 4.8 (a)]. It implies that this signature plot is a characteristic of domain 1 in PduB'[WT]. Deconvoluting this plot fit to two peaks giving transitions at 50.7±0.8^{o}C and 65.8±0.6^{o}C [Figure 4.8 (b)]. However, PduB'[D2] showed characteristics similar to PduA[WT] [Figure 4.8 (g,h)]. The calculated T_m for PduB'[D2] is 64.7 ± 0.07^{o}C from Boltzmann fit [Figure 4.8(c)] which is high in comparison to PduB'[D1]. Peak deconvolution fit to two peaks with 58.4 ± 0.8^{o}C and 71.0± 2.8^{o}C transition temperatures [Figure 4.8(d)]. Thus, PduB'[D2] is more stable than PduB'[D1]. On comparing its structurally equivalent protein generated from domain 2 of PduB', PduB'[CP-D2], it is observed to have comparable T_m to PduA[WT] i.e., 60.2±0.1^{o}C [Figure 4.8(e)]. It has also fit to two peaks with transition temperatures at 54.8±0.5^{o}C and 69.7±1.2^{o}C [Figure 4.8(f)]. So, overall it is seen that the PduB'[D2] has higher T_m than PduB'[D1] and its circular permutant variant has low temperature with an overall 4^{o}C difference. However, T_m of PduB'[CP-D2] and PduA[WT] are comparable and they both share the same topology.

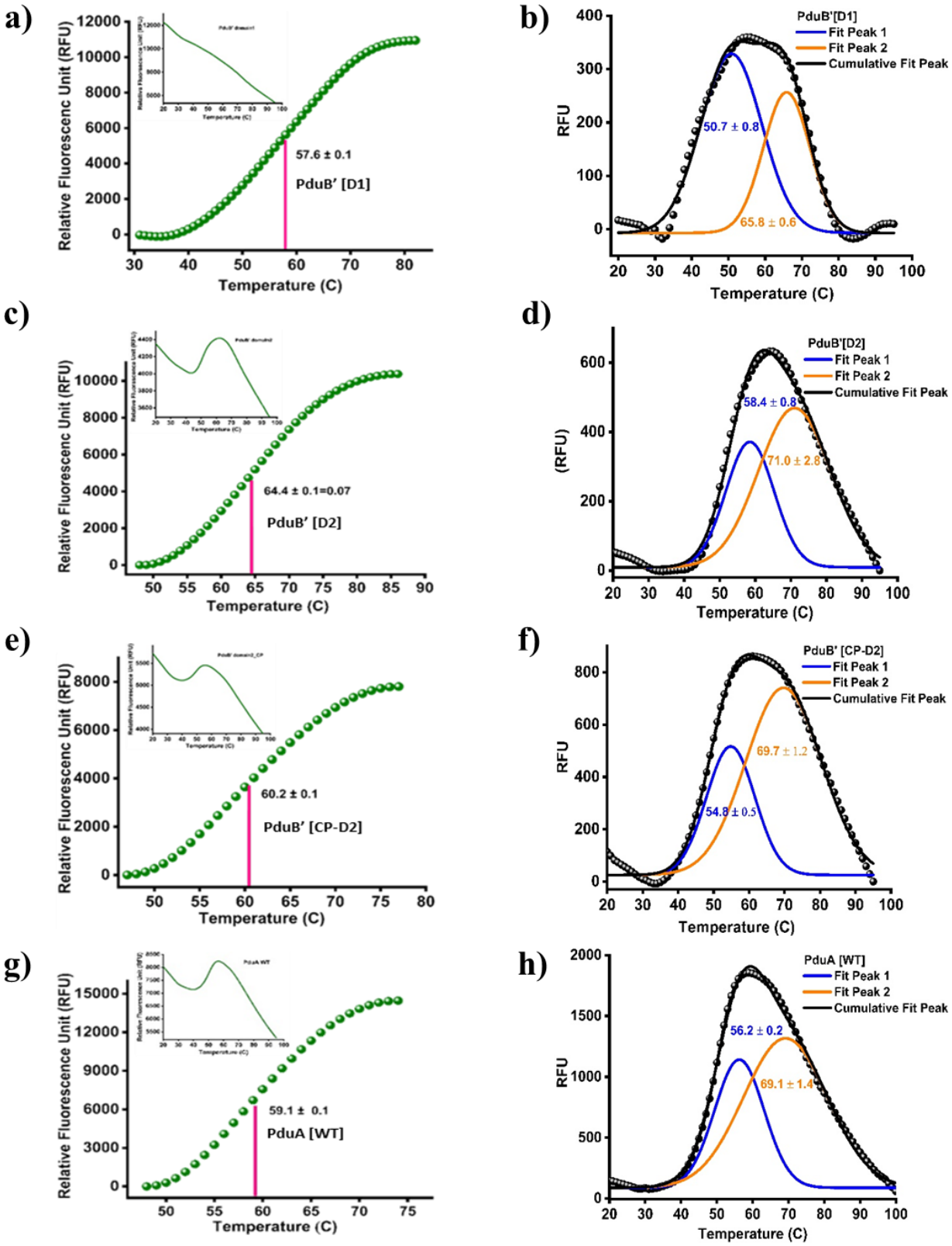

a)
Relative Fluorescenc Unit (RFU)
Temperature (C)
57.6 ± 0.1
PduB' [D1]
b)
PduB'[D1]
Fit Peak 1
Fit Peak 2
Cumulative Fit Peak
RFU
50.7 ± 0.8
65.8 ± 0.6
Temperature (C)
c)
Relative Fluorescenc Unit (RFU)
Temperature (C)
64.4 ± 0.1=0.07
PduB' [D2]
d)
PduB'[D2]
Fit Peak 1
Fit Peak 2
Cumulative Fit Peak
(RFU)
58.4 ± 0.8
71.0 ± 2.8
Temperature (C)
e)
Relative Fluorescenc Unit (RFU)
Temperature (C)
60.2 ± 0.1
PduB' [CP-D2]
f)
PduB' [CP-D2]
Fit Peak 1
Fit Peak 2
Cumulative Fit Peak
RFU
69.7 ± 1.2
54.8 ± 0.5
Temperature (C)
g)
Relative Fluorescence Unit (RFU)
Temperature (C)
59.1 ± 0.1
PduA [WT]
h)
PduA [WT]
Fit Peak 1
Fit Peak 2
Cumulative Fit Peak
RFU
56.2 ± 0.2
69.1 ± 1.4
Temperature (C)

Figure 4.8 Fluorescence-based temperature-dependent studies of single domain variants using Differential Scanning Fluorimetry (DSF). a) Boltzmann fit after baseline correction resulting in T_m 57.6 ± 0.1 ^oC for PduB'[D1]. Raw RFU curve (inset), b) DSF curve fitted to multiple transition peaks. PduB'[D1] fit to two transition T_m 50.7±0.8 ^oC and 65.8±0.6 ^oC; c) Boltzmann fit after baseline correction resulting in T_m 64.4 ± 0.07 ^oC for PduB'[D2]. Raw RFU curve (inset), d) DSF curve fitted to multiple transition peaks. PduB'[D2] fit to two transitions Tm 58.4 ± 0.8 ^oC and 71.0± 2.8 ^oC; e) Boltzmann fit after baseline correction resulting in T_m 60.2 ± 0.1 ^oC for PduB'[CP-D2]. Raw RFU curve (inset), f) DSF curve fitted to multiple transition peaks. PduB'[CP-D2] fit to two transitions T_m 54.8 ± 0.5 ^oC and 69.7±1.2 ^oC; g) Boltzmann fit after baseline correction resulting in T_m 59.1 ± 0.1 ^oC for PduA[WT]. Raw RFU curve (inset). h) DSF curve fitted to multiple transition peaks; PduA[WT] fit to two transition peaks at T_m 56.2 ± 0.2 ^oC and T_m 69.1 ± 0.14 ^oC. T_m values are marked by a red vertical line in Boltzmann fit. Transition temperatures are mentioned at the peak positions in multiple peak fit curves. DSF curves shown here are representations of three independent experiments.

Table 4.3 Transition temperatures for single domain protein variants obtained from Differential Scanning Fluorimetry, determined by Boltzmann fit and multiple peaks fitting and Circular Dichroism determined by Boltzmann fit.

	CD (Boltzmann fit)	DSF (Boltzmann fit)	DSF (Multiple peak fit)
PduA [WT]	-	59.1 ± 0.1°C	56.2 ±0.2 °C 69.1 ±1.4 °C
PduB'[D1]	-	57.6 C ± 0.1 °C	50.7±0.8 °C 65.8±0.6 °C
PduB'[D2]	-	64.7 C ± 0.07 °C	58.4 ± 0.8 °C 71.0± 2.8 °C
PduB'[CP-D2]	-	60.2 C ± 0.1 °C	54.8 ± 0.5 °C 69.7±1.2 °C

4.7 Conclusion

The evolution of complex microcompartment structures is a thoughtful and alluring question. The protein-based organelle has embedded a lot of diversity into it. The shell proteins are composed of single domain and tandem domains, self-assembling to form a compartment. Understanding the structure of MCP has become an utmost importance to understand the functional role and open more directions towards the generation of nanoreactors, nanostructures, and some targets hitting the pathogenicity of MCP. In this chapter, we tried to understand the prevalence of single domains in the MCP shell proteins. For this purpose, we dissected the tandem domain of PduB' [WT] into two domains individually, domain1 (PduB'[D1]) and domain2 (PduB'[D2]). PduA [WT] single-domain protein was observed to be a stable protein with respect to high temperatures (summarized in Table 4.3). On the other hand, PduB'[WT] showed low tolerance towards temperature and was found to be mobile as compared to highly compact PduA[WT]. So, we begin with generating PduB'[D1] and PduB'[D2] single-domain proteins. We also added another synthetic protein by

considering the 3D topology of proteins. We generated a protein equivalent to PduA[WT] topology from domain2 of PduB'[WT] sequence. The key findings from these studies are: We found PduB'[D1] to be responsible for a flexible or dynamic nature as indicated by high flexibility scores and high rmsf values. When studying temperature-dependent stability, we found single domains to be stable similar to PduA[WT]. Thus, the single domain variants are found to be more compact in structure and to provide BMC a flexible breathing shell, dimerization is a key. The presence of both domains together in PduB'[WT] allows its structure to be flexible and resilient towards a wide range of external stimuli. Also, we found that rearrangements in the structure of protein PduB'[D2] have important implications which led to reduced ability to self-assemble from extended sheets by PduB'[D2] to small sheets PduB'[CP-D2]. This property can be further explored to generate new different protein nanostructures.

CHAPTER – 5

In-vivo development and studies by gene substitution and gene swapping in Pdu-MCP operon

5.1 Introduction

Proteins emerge with distinct properties and functions at the level of divergence in amino acid sequences and modular rearrangements in the structural domains[116,117]. The evolutionary mechanisms underlying divergence are duplication, deletion, fission, fusion, and recombination events[118,119]. Even conserved families of proteins have rearrangements and reoccur in different orders leading to circularly permuted proteins[102]. The Pdu microcompartment has a fascinating architecture with eight different shell proteins making the covering for the encapsulated enzymes. Eight different shell proteins poses different sequence and structures though belonging to a conserved family of bacterial microcompartment domain family. This diversity in the shell proteins is intriguing. The answer to the question of role played by two types of oligomerization of shell proteins as hexamer and pseudohexamer for making the shell will ease the development of nanoreactors or nanocarriers in future. In the previous two chapters, we had attempted to address this question using biophysical approaches *in vitro*, where we strategically developed dimer-domain and single-domain synthetic protein variants. We found single domains proteins to be stable and dimer proteins has a role in providing flexibility to the microcompartment.

In this chapter, we explored one of the major shell proteins that exist as a single-domain to be substituted with its structurally equivalent synthetic protein variant in an operon *in vivo*. Some of the bacterial microcompartment shell proteins reveal rearrangements as circular permutations where a similar 3D structure is built from secondary structure elements containing new N and C terminal regions. BMC shell proteins derived from the conserved domain family of BMC, oligomerizing into a higher-order similar structures, have diverse functions. Interestingly, these shell proteins are clustered together and transcribed from a single operon[39] indicating a highly regulated system. About 50% of the prokaryote genes exist in operons[120] and it is well established that co-expression of genes involved in the same metabolic pathway are often encoded on an operon and arranged in a particular order[121]. In one of the reports, the importance of gene location in an operon was explored by focusing on PduJ and PduA which are major shell proteins of Pdu MCP and have 80% identical amino acid sequences[64]. On chromosomally replacing the PduA gene with PduJ, the protein could mediate the function of PduA of 1,2 propanediol transport into the MCP. This suggests that protein interactions with adjacent proteins in MCP higher-order assembly

are governed by gene position or gene order. PduA is a major shell protein and occupies the first gene position on operon and has been reported to mediate maximum interactions with other proteins[55,63]. This interesting data encouraged us to study the significance of gene positioning in Pdu MCP operon mediating protein-protein interactions. The new circularly permuted single-domain variant generated in Chapter 4 with native sequence when replaced with their native counterpart in an operon will provide insights into the evolutionary role of diverse sequences and structure of shell proteins and its implications in the global stability of BMC. Therefore, in this chapter, we have developed *in vivo* mutant strains where the first two positions of the Pdu operon have been explored. First, we developed a strain where we incorporated the protein variant in place of PduA *invivo.* PduA is a well-studied protein that has its solved crystal structure[42]. Replacing PduA with its circular permutant variant *in vivo* which is similar in topology will be an interesting study. We have developed this variant by choosing the sequence of the second domain of PduB' which was circularly permuted to obtain a structurally similar protein like native PduA (ΔPduA: PduB'[CP-D2]). Furthermore, the strain was developed where PduA gene position is switched by the second gene PduB'. So a new variant is generated which has PduB' gene in first position and PduA occupies the second position. This new operon will give insight into the positional role of genes and may result in varying protein-protein interactions while biogenesis of MCP. Schematics to generate the new strains is given in [Figure 5.1].

Previously, to understand the role of a particular protein in microcompartment assembly and function, many studies have been conducted where individual genes were deleted and their impact was studied[31,65,71,73,111]. Understanding the microcompartment architecture and assembly principles is valuable which will help in designing new structures or compartments to be developed as nanoreactors or delivery cargos. A report has shown where a circularly permutant protein was generated from PduA protein and a 13 nm dodecahedral cage was obtained[122]. In another study, a circularly permuted PduA was developed to explore the localization of proteins towards the surface or lumen of the BMCs[123]. BMC-H protein from *Haliangium ochraceum (*HO)shell was also explored for precise control over quantitative cargo incorporation by its circularly permutant version[124]. However, circular permutations in the context of understanding the microcompartment structure and function relationship were still elusive.

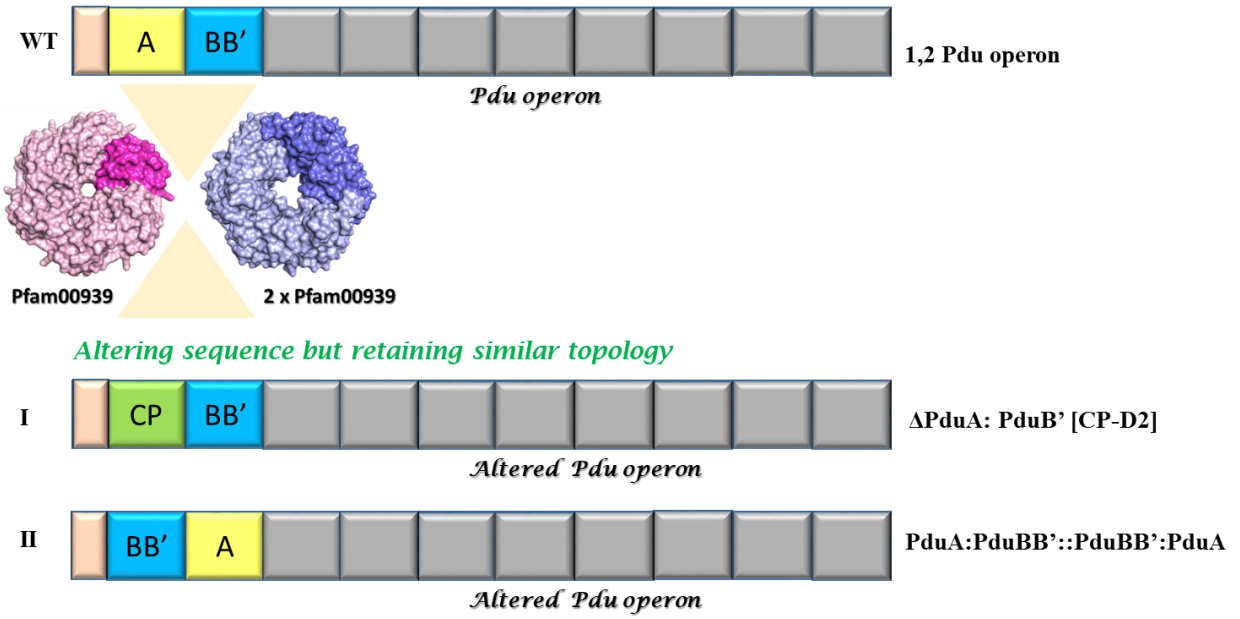


Figure 5.1 Schematics showing the theme of *invivo* Pdu operon alteration. Pdu operon showing only first two genes PduA, PduBB' (I) Pdu operon is altered at the first gene for PduA, where it is replaced by a circular permutant PduB'domain2 protein sequence [ΔPduA: PduB' [CP-D2]]. (II) The first and second gene positions are swapped from PduA at first place and PduBB' at second place to PduBB' being at first gene position followed by PduA at second position in an altered swapped operon [PduA:PduBB'::PduBB':PduA].

5.2 Methodology to develop the operon-modified strains using CRISPR-Cas9 gene-editing tool

For the development of operon-modified strains, firstly, lab strain *Salmonella enterica* serovar Typhimurium LT2 (BE293) which carries pkD46 plasmid was cured. This plasmid carries the ampicillin resistance gene. It was cured by repeatedly growing the strain in the absence of the antibiotic, till it gained the sensitivity towards antibiotic. This cured mutant (SSL1) was employed for genome editing studies using CRISPR Cas9 Technology. Initially, plasmid pCasM was

electroporated into *Salmonella enteria* LT2 (SSL1) which contains both Cas9 and a λ-Red recombineering system[93]. Electroporation was performed in a 2-mm cuvette at 2.5 kV, 200 Ω, and immediately 1ml ice-cold SOC media was added to resuspend the cells and incubated at 37 ^{0}C. The cells were allowed to recover for 2 hr at 37 ^{0}C. The transformants were then selected on a Kanamycin antibiotic selection plate and used to prepare electrocompetent cells for the insertion of an antibiotic resistance cassette. The electrocompetent cells were prepared at ~0.6 $OD_{600.}$ The harvested cells were then washed three times and finally resuspended in ice-chilled water containing 10% glycerol. The cells were concentrated to 50 µl volume and used for electroporation.

The antibiotic resistance cassette, used here is ampicillin (bla) resistance cassette (ARC) and was prepared by overlap extension PCR to be inserted at the target site, PduA gene is taken as an example. The ampicillin resistance gene (bla) was PCR amplified and fused with upstream and downstream homologous arms of the target site. The generated ARC cassette was electroporated into the above-prepared electrocompetent cells (50 µl) at a concentration of 300 ng. Transformants were screened using an agar medium containing Kanamycin (50 µg/ml) and Ampicillin (100 µg/ml) antibiotics. The colonies were screened by PCR screening method and sequencing after genomic DNA isolation. Positive clones were then further used to prepare electrocompetent cells to transform the desired gene (donor DNA). The desired gene to be inserted was also generated by the overlap extension PCR method containing the upstream and downstream homology arms. Donor DNA and psgRNA were electroporated at 300 ng and 200 ng concentrations to the above prepared electrocompetent cells respectively. Transformants were screened on Kanamycin (50 µg/ml) and spectinomycin (200 µg/ml) antibiotics. Further positive modified clones were confirmed using the PCR screening method and Sanger sequencing. The general scheme of the methodology is described in [Figure 5.2].

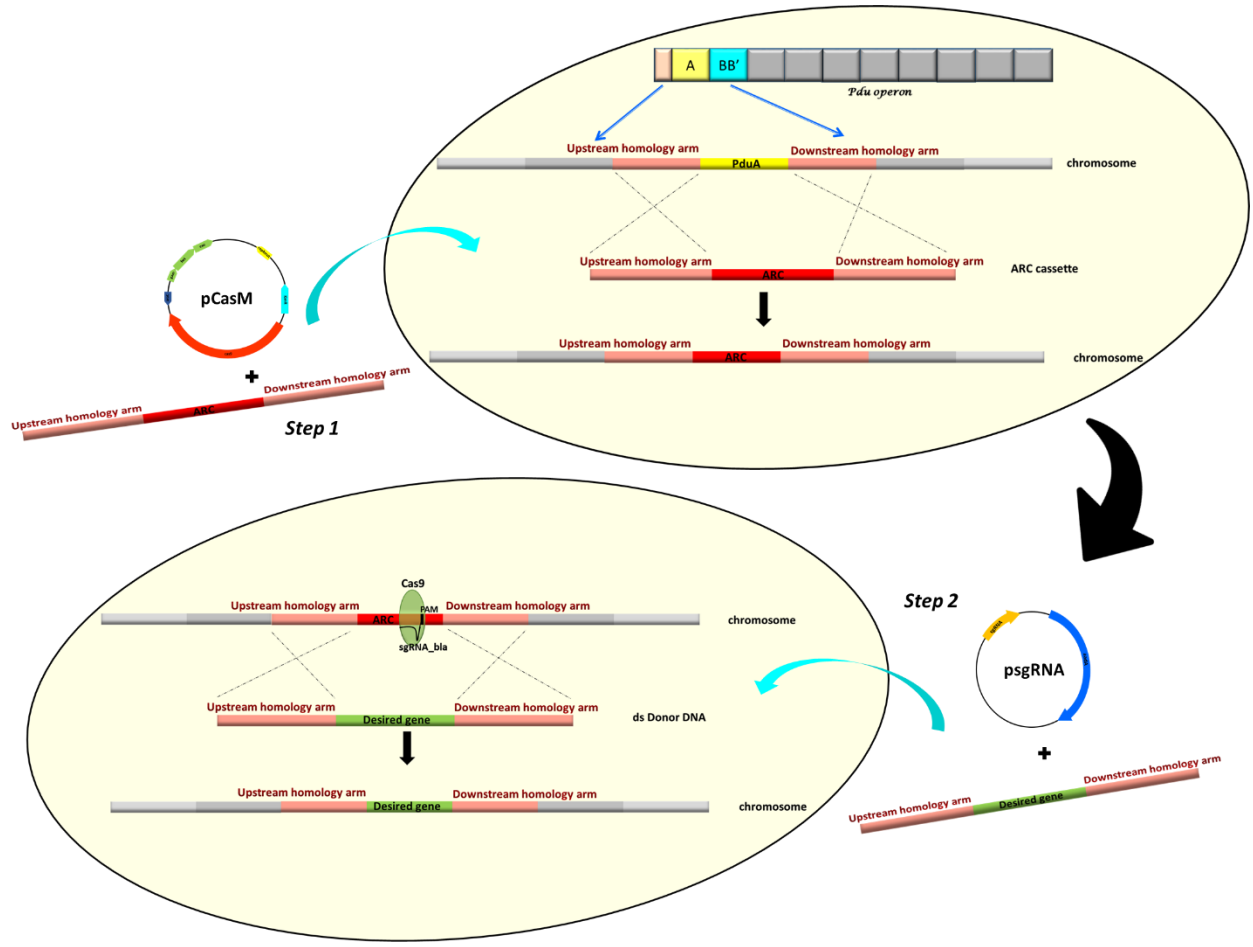


Figure 5.2 Methodology for the development of operon-modified strains. PduA gene (yellow) was taken as an example. Step 1 includes the electroporation of an antibiotic resistance cassette (red) with homology arms (ARC-bla) into the pcasM harboring cells where the target gene, PduA will be replaced by ARC. Positive transformants are guided by ampicillin selection. Step 2 then employs the electroporation of psgRNA and the donor template (green) with homology arms. Here sgRNA coupled with Cas9 will target and create a dsDNA break at bla gene and a donor template will be inserted as a result of homologous recombination. Positive transformants are selected on spectinomycin and kanamycin selection. They are further verified using sequencing.

5.3 Development of strain containing modified pdu operon, ΔPduA:PduB'[CP-D2]

PduA gene deletion studies conducted earlier have shown that strains form enlarged MCPs as compared to wild type and growth studies have been shown to induce a growth arrest period for 15-20h at saturating B12 concentration. Thus, suggesting that PduA is an essential component for MCP formation and alleviating propionaldehyde toxicity. We have developed a strain that has been modified to carry circularly permuted PduB' domain2 in place of PduA. The second domain of PduB' was chosen based on alignment scores, where it is better aligned with PduA with RMSD 0.85Å than domain 1 with RMSD 1.49Å (mentioned in Chapter 4). A circular permutant was generated by identifying the linker and terminal ends of both proteins and subsequently modifying the PduB' domain2 according to the PduA structure. Schematics of the operon is shown in [Figure 5.3] where PduA gene is highlighted with upstream and downstream homology arms and a modified operon with the CP variant in place of PduA. As explained in the general scheme [Figure 5.2] for the development of modified strains, an antibiotic resistance cassette (bla) was first generated using overlap extension PCR with the upstream and downstream homology arms of PduA i.e., PduF and PduBB' respectively. This ARC was used to electroporate in BE293 cured strain, SSL1, containing pCasM plasmid. The positive transformants harboring ARC replacing PduA gene were then selected on kanamycin and ampicillin selection agar plates followed by PCR screening methodology and Sanger sequencing. The confirmed strain was then used to make electrocompetent cells for the transformation of the donor template and psgRNA. The donor template was generated by the overlap extension PCR method where the desired gene sequence (PduB'[CP-D2]) to be replaced at PduA is amplified and the template for this PCR is used from the constructs generated in Chapter 4. The donor amplicon was fused with upstream and downstream homology arms. psgRNA and donor template were co-transformed into the strain harboring ARC (SSL2). Cas9 coupled with sgRNA will create a dsDNA break at bla gene and the presence of a donor template will allow homologous recombination to occur at a specific site and a modified strain is produced. The screening for the desired positive strain is done on LB agar plates containing Spectinomycin and kanamycin antibiotics, followed by colony PCR screening using the primers for upstream and downstream homology arms. The agarose gel image shows the presence of the desired size of the amplicon that was inserted [Figure 5.3 (b)]. Further, the samples

were confirmed by the Sanger sequencing methodology for the presence of the ΔPduA: PduB'[CP-D2] gene as highlighted in blue [Figure 5.3 (c)]. The modified strain PduB'[CP-D2] substituted for PduA (SSL3) might result in different structure and functional properties of the microcompartment. However, comparing the biophysical properties of CP PduB'dom2 has shown it has similar temperature-dependent stability and self-assembles to form sheets, although smaller than PduA.

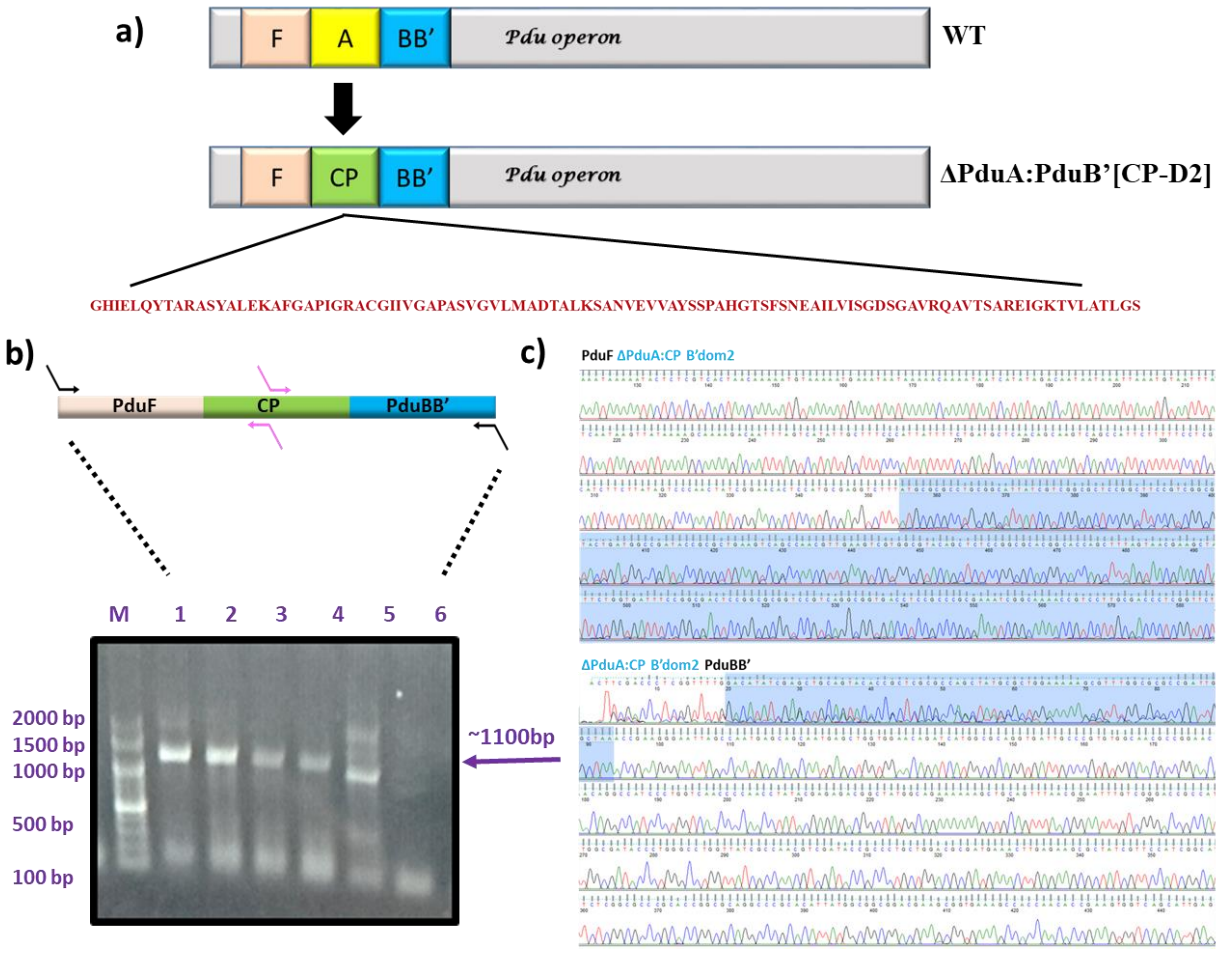


Figure 5.3 Development of modified Pdu operon ΔPduA: PduB'[CP-D2] a) Schematics showing the pdu operon with highlighted PduA gene (yellow) to be replaced along with upstream (PduF, orange) and downstream (PduBB', blue) homology arms as a WT operon (top). Modified pdu operon with replaced circularly permuted PduB'domain 2(green), mentioned as CP (bottom). Sequence of PduB' [CP-D2] is shown in red. b) Agarose gel showing the screening of positive strains. Lanes 1-6 are different colonies obtained and screened. Screening was done using the primers for upstream and downstream homology arms. C) Sequencing of the positive strain showing successful replacement of PduA gene with PduB' [CP-D2] (highlighted in blue).

5.4 Development of strain containing modified pdu operon PduA:PduBB'::PduBB':PduA

We have developed one strain where we substituted PduA at the first position to its structurally similar variant generated from PduB' domain2. Now we are further interested in elucidating the role of gene position. PduA being the first protein in an operon is believed to act as an initial and important support pillar for the interaction and assembly of the microcompartment. It has been seen previously that the gene order is important for the biogenesis in MCP. Previously, the role of PduJ which is 80% identical to PduA, and both oligomerizing as hexamers was studied by repositioning both the genes. Here, we are exploring the first two positions in pdu operon to reposition and swap the genes. The first position is occupied by PduA which oligomerizes as a hexamer and the second position is occupied by PduBB' which is a pseudohexamer. The gene swapping of a hexamer and a pseudohexamer will provide insights into the structural and positional role of these proteins. So, to generate a strain having PduA and PduBB' swapped genes, we first amplified the ARC cassette (bla) containing upstream and downstream homology arms i.e., PduF and PduC using overlap extension PCR. This ARC was then inserted into pCasM harboring strain as done for the previous strain. The positive strains are selected on LB agar plates containing ampicillin antibiotics and verified using PCR screening and Sanger sequencing methodology. The confirmed strain was then used to prepare electrocompetent cells. Donor

template containing PduBB' fused to PduA in 5'to 3' direction was PCR amplified keeping PduBB' upstream to PduA using overlap extension PCR with the same homology arms. The donor template was electroporated into prepared electrocompetent cells and screened on spectinomycin and kanamycin antibiotic plates. Further, obtained colonies are screened by PCR amplification using the combination of primers as shown in the agarose gel image [Figure 5.4]. The positive colonies containing the modified operon [PduA:PduBB'::PduBB':PduA] strain (SSL4) were verified by Sanger sequencing.

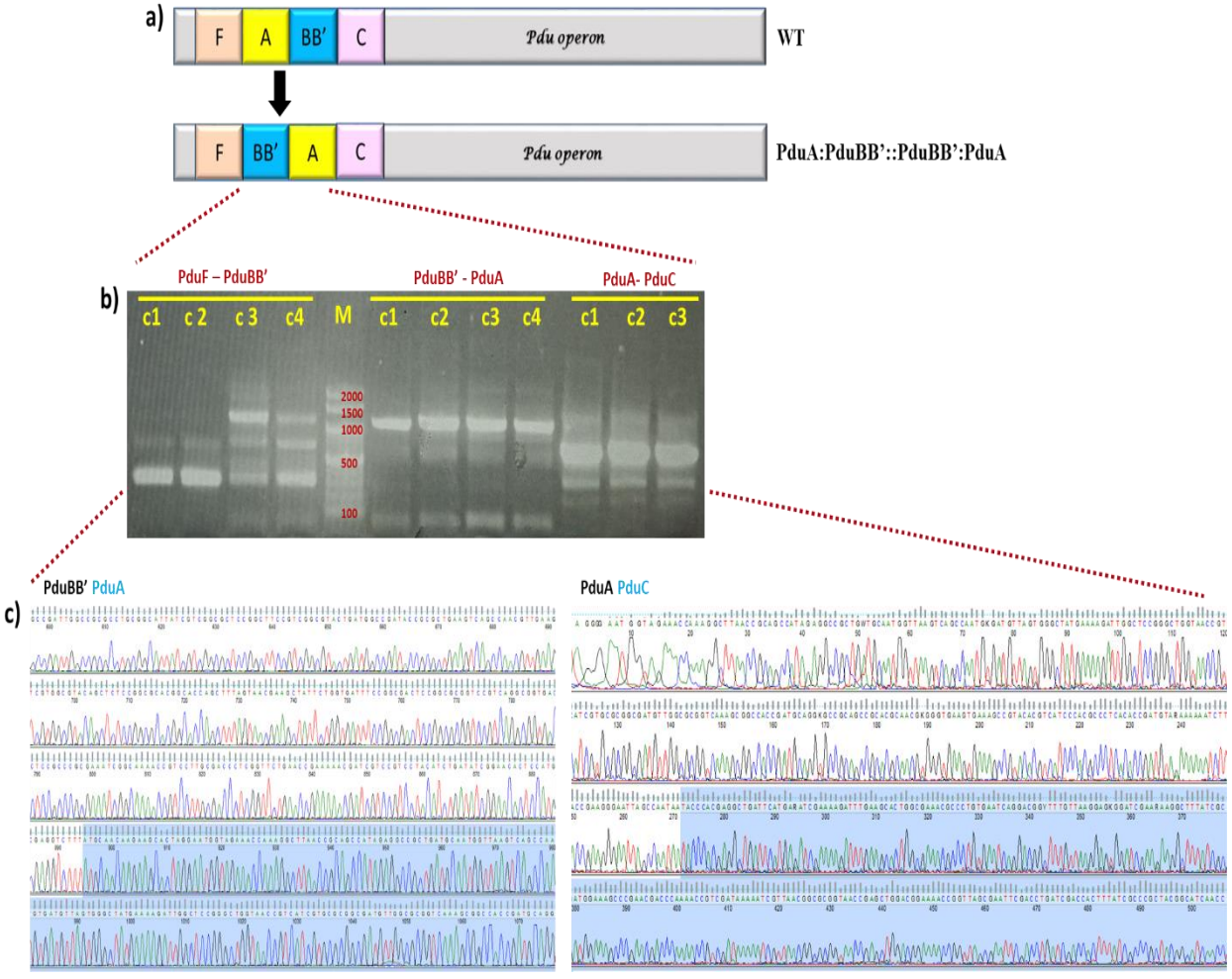


Figure 5.4 Development of modified pdu operon PduA:PduBB'::PduBB':PduA a) Schematics showing the pdu operon with highlighted PduA gene (yellow) and PduBB' (blue) to be swapped along with upstream (PduF, orange) and downstream (PduC, pink) homology arms as a WT operon (top). Modified pdu operon with swapped PduA and PduBB'. PduBB' acquired the first position and PduA second position in an operon (bottom). b) Agarose gel showing the screening of positive strains. C1-C4 are four different colonies obtained. Screening was done using different primer combinations of gene-specific and upstream and downstream homology arms. C) Sequencing of the positive strain showing successful gene swapping [PduA:PduBB'::PduBB':PduA].

5.5 Growth curve analysis of modified strains.

After the sequence of modified strains was confirmed, further, to determine whether the shell protein replacements have a substantial effect on the growth of the bacteria, we evaluated the growth curve of wild-type and modified strains in minimal media in the presence of an inducer, 0.6% 1,2 propanediol (1,2 PD) supplemented with 150 nM CN-B_{12} cofactor as reported earlier for growth assays. We found a longer lag phase and delayed growth of strains than wild type. The modified strain, ΔPduA: PduB'[CP-D2] grew slower than the wild-type strain. We observed a long lag phase which persisted till 30 30-hour period before the onset of growth [Figure 5.5 (a)]. The delayed lag phase and slow growth suggest toxicity due to the intermediate propionaldehyde on metabolizing 1,2 propanediol for its growth when PduA is substituted to its structurally equivalent variant PduB' [CP-D2] having a different sequence. This further implies improper MCP formation. This result is different from previous studies where deletion of PduA resulted in growth arrest for a shorter period of time and enlarged MCP. On analyzing strain another strain PduA:PduBB'::PduBB':PduA, where the first two proteins with different oligomerization are swapped, we found a growth arrest period for 10-12 hr initiated from 8h which persisted till 20hr [Figure 5.5 (b)]. This strain later reached a similar growth density as a wild-type strain rescuing the growth of the bacteria. These results are similar to previous results where PduA is deleted and growth arrest happens initially and later growth is supported. A slow growth initially indicates

leaky or defective MCP leading to propanaldehyde toxicity. This result indicates that even after swapping the first two important proteins the bacteria can mitigate propanediol toxicity.

Further, to assess the impact of substituting and swapping genes in MCP operon we tried to purify the microcompartment from the modified strains. To our surprise and after many attempts, we could not purify the microcompartment. Thus, the substitution of PduA[WT] with its structurally similar protein variant could not replace it. This can be defined by the fact that the PduB'[CP-D2] when purified and studied *in vitro* could not form long extended sheets as PduA. This suggests the important role of PduA sequence in forming the extended sheets as facets of MCP and in keeping the necessary interactions with other proteins

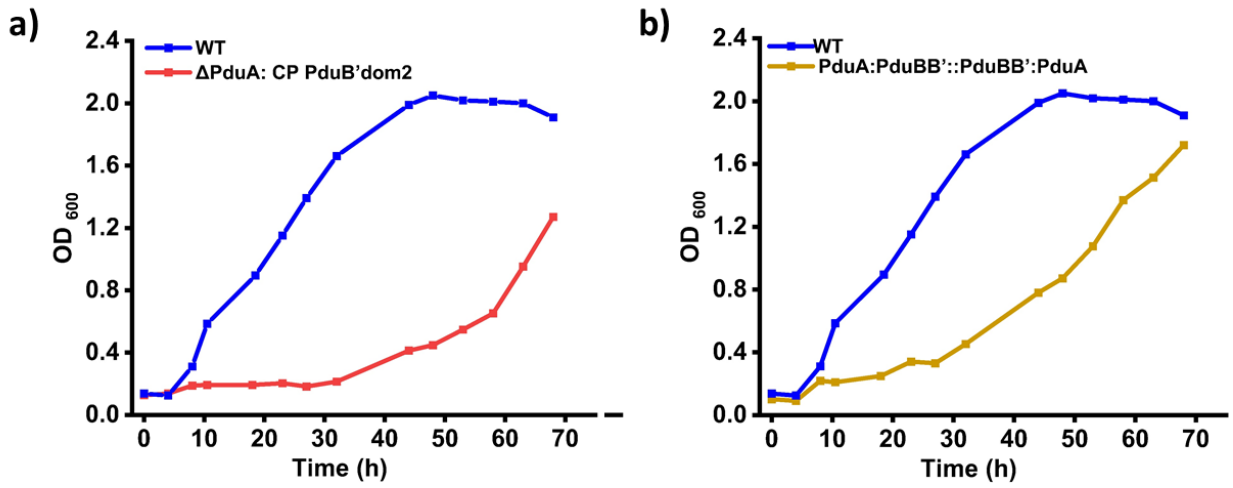


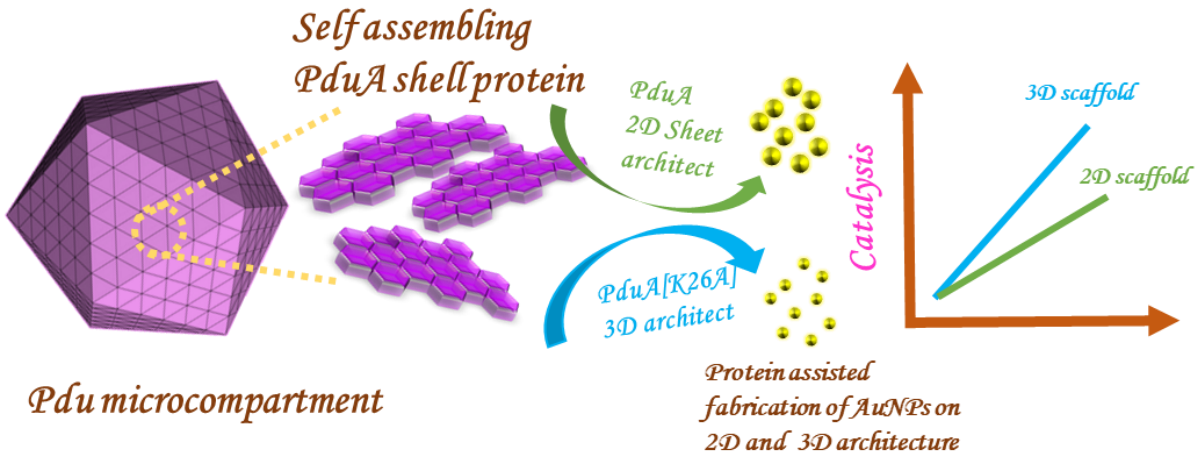
Figure 5.5 Growth curve of an operon modified strain. a) ΔPduA: PduB' [CP-D2] b) PduA: PduBB':PduBB': PduA. Growth curves were generated by allowing modified strains to grow on a minimal media with 0.6% 1,2 propanediol and 150 nM CN-B_{12}.

5.6 Conclusion

The present work focuses our attention and allows us to have some insights into the importance of shell protein sequence and structure in the microcompartment architecture. MCP architecture is designed by nature in a way to protect the cell from toxic and volatile intermediates in the metabolic pathways of certain substrates. The presence of different shell proteins with different oligomerizations has important structural and functional implications. In an attempt to understand the structural role and sequence effect on MCP architecture and stability, we substituted PduA with PduB' [CP-D2]. The positional role of a hexamer and a pseudohexamer was explored by substituting their positions and a mutant strain, PduA: PduBB':PduBB': PduA was developed. It is worth noticing and one can wonder why a hexamer is placed in the first position in an operon. What are the implications of substituting that protein with a protein of different oligomerization in the first place in an operon? Does the oligomerization state or the sequence of the protein plays an important role? In this work, we observed delayed growth rates and could not purify the microcompartment from the modified strains. Further, more insights can be delivered from more experiments and critical analysis in addition to these preliminary studies presented in this thesis chapter. In the future, we can explore more of the protein variants by chromosomal substitutions in an operon with their native counterparts. This will aid in further understanding the structural roles of shell protein paralogs as a dimer or single domains and their exquisite sequence specificities.

CHAPTER – 6

BMC shell proteins as scaffolds in varying architectures for modulating properties of catalytic gold nanoparticles



6.1 Introduction

In the previous chapters, we have explored two major shell proteins to understand the implications of the modularity of shell proteins focusing on the sequence and structure of shell proteins. In that search, we designed a few protein variants (Chapters 3 and 4) and also attempted to study them *invivo* while considering one of the variants (Chapter 5). Further, we also studied the structural and positional role of the two essential shell proteins, PduA (hexamer) and PduBB' (pseudohexamer) occupying the first and second positions in an operon respectively (Chapter 5). From the above studies, we found native shell proteins and synthetic protein variants to be thermostable and PduA shell protein is found to be a more rigid and compact protein as compared to other shell proteins. So, we believe this protein can be explored as a stable scaffold for decorating nanoparticles, enzymes, etc for catalytic applications. One advantage of using this shell protein is that it has a solved crystal structure and it is a naturally self-assembling protein synthesized as 2D sheet. So, we attempted to fabricate nanoparticles on the 2D sheet and study their catalytic function. The function of nanoparticles not only depends on the properties of the nanoparticles but also their spatial orientations. Varying architectural designs for scaffolds can be used to develop different orientations of nanoparticles in three dimensions. Keeping that in mind, we employed a single-point mutant of PduA, (PduA[K26A]) for the fabrication of gold nanoparticles in 3D architectures[125]. This approach will demonstrate the hierarchical organization of the scaffold can assist in the fine-tuning of nanoparticle properties, thus paving the way for the design of new nanoscale materials.

Metal nanoparticle synthesis has attracted substantial interest and acceptance in the area of bio-diagnostics, sensing, bio-imaging, and catalysis[126–132]. The versatility of nanoparticles lies largely in their different shapes and sizes[133–136]. In comparison to their bulk counterparts, nanoparticles offer a much higher surface-to-volume ratio and therefore exhibit an excellent catalytic activity. Heterogeneous catalysis which occurs at the interface between a solid catalyst nanoparticle and liquid/gaseous reactants is dependent on nanoparticle shape and size. It determines the availability of active sites on the nanoparticle surface, affecting the binding strength of reactants and intermediates[137,138]. There have been many reports of nanoparticle synthesis using various physio-chemical methods[139–146]. Biomolecule-assisted synthesis and fabrication of catalytic metal

nanoparticles have of late attracted a lot of attention in the field owing to the availability of versatile substrates, mild reaction conditions, and ease of manipulation. Proteins serve as a potential scaffold for reducing metal ions and nanoparticle fabrication and also prevent agglomeration by restraining particle size growth[147–154]. Further, proteins may be made into various 2-dimensional or 3-dimensional scaffolds by simple genetic engineering procedures making them very attractive materials for the development of organic–inorganic hybrids[129,155]. Specific arrangements of the nanoparticles also can be attained by meticulous protein design and smart positioning of specific amino acids that bind to the nanoparticles. Various reports have shown that certain amino acids, such as lysine, aspartic acid, glutamic acid, tyrosine, tryptophan, and cysteine in proteins, are involved in the reduction of gold precursors to gold nanoparticles[156–158]. Hence, the use of proteins that can form different scaffolds in two or three dimensions can be used to fabricate different arrangements of nanoparticles leading to altered properties. To date, the literature has mostly explored the globular proteins or small peptides in fabricating nanoparticles[148,149,159–163], and proteins forming supra-molecular assemblies in other architectures remain underexplored.

For this study, two proteins are used which form supramolecular assemblies naturally on expression and purification and are variants of PduA protein. Reports have shown that a point mutation in PduA leads to the loss of the polyhedral structure of MCP, along with the abolishment of its sheet structure, and leads to the formation of an aberrant 3D non-sheet morphology[125,164]. The crystal structure of PduA reveals the formation of stable well-ordered 2D protein sheets where hexamers are involved in edge-to-edge interactions [Figure 6.1(a–c)]. Earlier reports by Sinha et al. (2014), have shown that the lysine at the 26th position is crucial for the interactions in forming extended 2D sheets and mutating the lysine residue to alanine (K26A) in PduA alters the edge–edge interaction to edge–face interactions, here shown in [Figure 6.1(d, e)] giving rise to an aberrant 3D architecture to a mutant protein, PduA[K26A] [Figure 6.1(f)]. Its solved crystal structure revealed two crystal forms that assemble to form supramolecular 3D architecture. Crystal form 1 reveals the three hexamers contact edge to face and crystal form 2 reveals the two interacting edges form extended one-dimensional strips where edges interact with faces of other hexamers. Earlier, the shell protein surface of intact MCP has been shown to reduce gold salt precursor to gold nanoparticles and developed an organic–inorganic hybrid catalyst[165]. Here, we

have used the major shell protein, PduA (PDB ID 3NGK)[166] of Pdu MCP, and its mutant PduA[K26A] (PDB ID 4QIF)[125] as a natural biological scaffold for nanoparticle fabrication where both self-assembles into different architectures. Since it is the same protein with just a single mutation, the study is best suited for comparing the effects of shape and dimensions of supramolecular organic assemblies in fabricating metal nanoparticles and studying their properties. We explored the effect of 2D and 3D scaffolds on the synthesis and catalytic properties of gold nanoparticles.

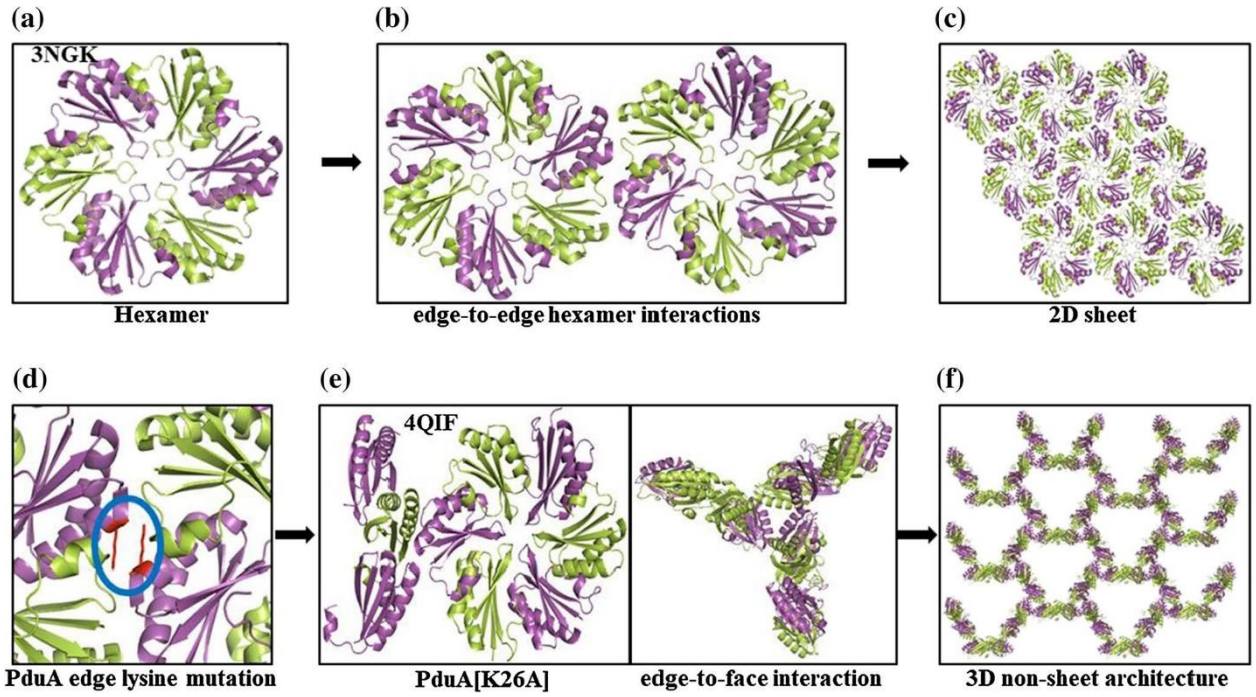


Figure 6.1 Crystal structure of protein scaffolds a) hexameric PduA scaffold protein (PDB-3NGK) b) which tiles edge-to-edge to form c) extended protein sheets in 2D; d) PduA protein showing lysine involved in edge-to-edge interactions which on mutation to alanine (K26A) disrupt the edge-to-edge interaction and e) hexamers contacts edge-to-face (PDB-4QIF) to form f) non-sheet 3D higher-order supramolecular structure molecular structure. Each alternative monomer unit is shown in different colors for better visual interpretation.

6.2 Gold nanoparticle fabrication on the 2D and 3D shell protein architecture

The effect of 2D and 3D scaffolds on the fabrication of gold nanoparticles is explored by using PduA and PduA[K26A] shell proteins. In vitro overexpression of PduA protein self-assembles as an extended 2D sheet and PduA[K26A] protein as an aberrant 3D architecture due to its disrupted edge interaction, (PDB 3NGK, 4QIF respectively). The overexpressed proteins' purity is confirmed on SDS-PAGE [Figure 6.2]. Unlike globular proteins, which require chemical agents or genetic engineering methods to crosslink monomer units to form 2D scaffolds, these proteins naturally self-assemble to form different spatial nanostructures. The uniform repetitive building blocks will allow precise and uniform arrangement of nanoparticles on their surface. Moreover, surface-exposed residues with wide charge distribution facilitate the interaction of metal salt precursors and their subsequent reduction into nanoparticles. To fabricate gold nanoparticles on these protein scaffolds, both proteins with two different architectures are incubated with gold chloride ($AuCl_3$) salt to allow the interaction with surface residues. This interaction results in the reduction of metal oxidation state (Au^{3+} to Au^0) by transfer of electrons from amino acids, such as lysine, arginine, tyrosine[152,156]. The reduced gold here acts as a nucleation site for the further growth of nanoparticles. Transmission Electron Microscope (TEM) images are analyzed after 12 h incubation along with the protein controls where we did not add $AuCl_3$ salt [Figure 6.3(a, d)]. PduA is seen as a sheet in control and PduA[K26A] as an aberrant structure. Uniformly distributed gold nanoparticles are seen where $AuCl_3$ salt is added [Fig. 6.3(b,c,e,f)]. TEM micrographs confirm the formation of gold nanoparticles on both the scaffolds. These results show that the shell surface residues facilitate the reduction of $AuCl_3$ to gold nanoparticles.

6.3 Spectroscopic studies of fabricated gold NPs on different shell protein scaffolds

Gold nanoparticles exhibit a strong absorption spectrum in the visible region that corresponds to the excitation of surface plasmon vibrations (surface plasmon resonance) of the conduction electrons on the metal surface[167,168]. The SPR peak for the gold nanoparticles is observed in the 500–600 nm wavelength range. In our experiments, we monitored the emergence of this peak for

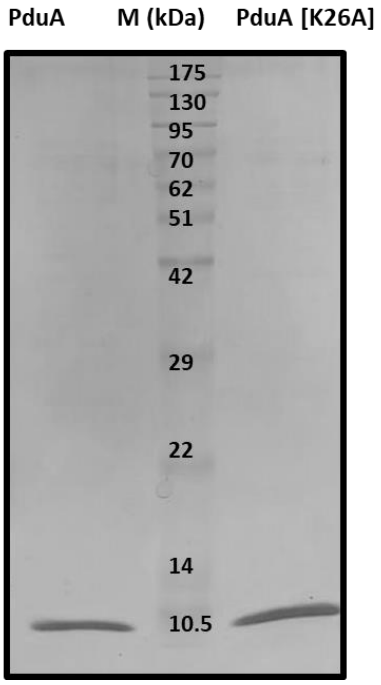


Figure 6.2 SDS PAGE of purified proteins.

measuring the kinetics of gold nanoparticle formation on the protein surface using a UV–vis spectrophotometer. As a control, bare protein and buffer are also monitored in this range where we did not observe any reduction in absorption spectra. UV–vis spectra observed on the reduction of gold nanoparticles with PduA 2D scaffold, showed an absorption peak at a wavelength of 575 nm, while in the case of the PduA[K26A] 3D scaffold, the absorption peak is observed at 555 nm wavelength. [Figure 6.4(a)]. This 20 nm shift in the wavelength corresponds to the differences in the size of nanoparticles formed[169,170]. When gold nanoparticles are reduced on the PduA scaffold, a redshift in wavelength indicates gold nanoparticles formed by PduA are large as compared to PduA[K26A] nanoparticles. We also observed a shorter lag time with PduA than the PduA[K26A] scaffold [Figure 6.4(b)]. This is probably due to the scaffold provided by PduA 2D sheets having easily accessible surface residues on both sides of the sheet which facilitate the rapid interaction of ionic gold precursor in solution to form nucleating sites, then subsequent reduction followed by

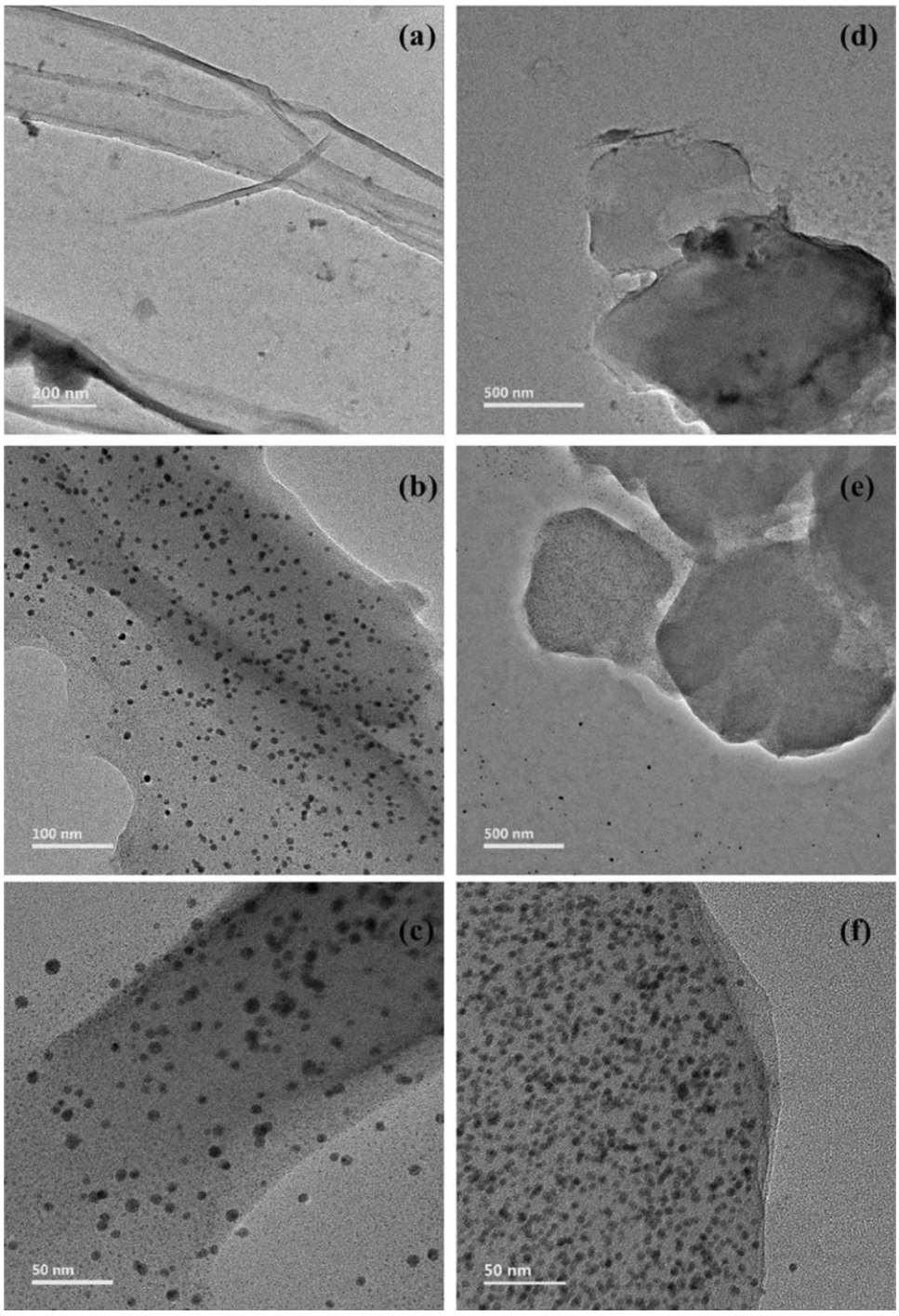


Figure 6.3 Transmission electron microscopy (TEM) images of protein scaffolds alone and with fabricated gold nanoparticles. (a) PduA 2D sheet (b) fabricated gold NPs on protein surface PduA 2D. Scale bar 100 nm, (c) 50 nm; (d) PduA[K26A] non-sheet morphology (e) fabricated gold nanoparticles on PduA[K26A] 3D. Scale bar 500 nm, (f) 50 nm.

growth of gold nanoparticles. On the other side, PduA[K26A] mutant with a large surface area due to 3D architecture has more interacting residues and thus forms more nucleating sites with gold precursors which results in a delayed reduction and small size of gold nanoparticles. This correlates with the theoretical solvent accessible surface area (SASA) calculated by Pymol[171] which predicts a higher surface area 74,207 $Å^2$ for PduA[K26A], structure generated at 54mer assembly (minimum self-assembled supramolecular structure) than PduA, with the surface area of 60,071.445 $Å^2$ at the same 54-mer self-assembled structure. The reduction of $AuCl_3$ to gold nanoparticles is attained via an intermediate state of gold nanocluster which is followed by their assembly and growth into polycrystalline gold nanoparticles[142,172]. Gold nanoclusters (AuNCs) with a dimension less than 2 nm exhibit fluorescence properties. Within such dimensions of a few atoms, the band structure breaks into discrete energy levels which leads to fluorescence in AuNCs[173]. To understand this mechanism, we measured the fluorescence of gold nanoclusters (AuNCs) formation before the nanoparticle stage by exciting at a wavelength of 370 nm. The increase in fluorescence intensity with time is observed with the maximum emission wavelength at 658 nm. [Figure. 6.4(c, d)]. Both the scaffold nanoparticles go through a similar nanocluster intermediate stage and interestingly result in different-sized nanoparticles. This suggests that the desired different properties of nanoparticles can be attained using the different structures of protein scaffolds.

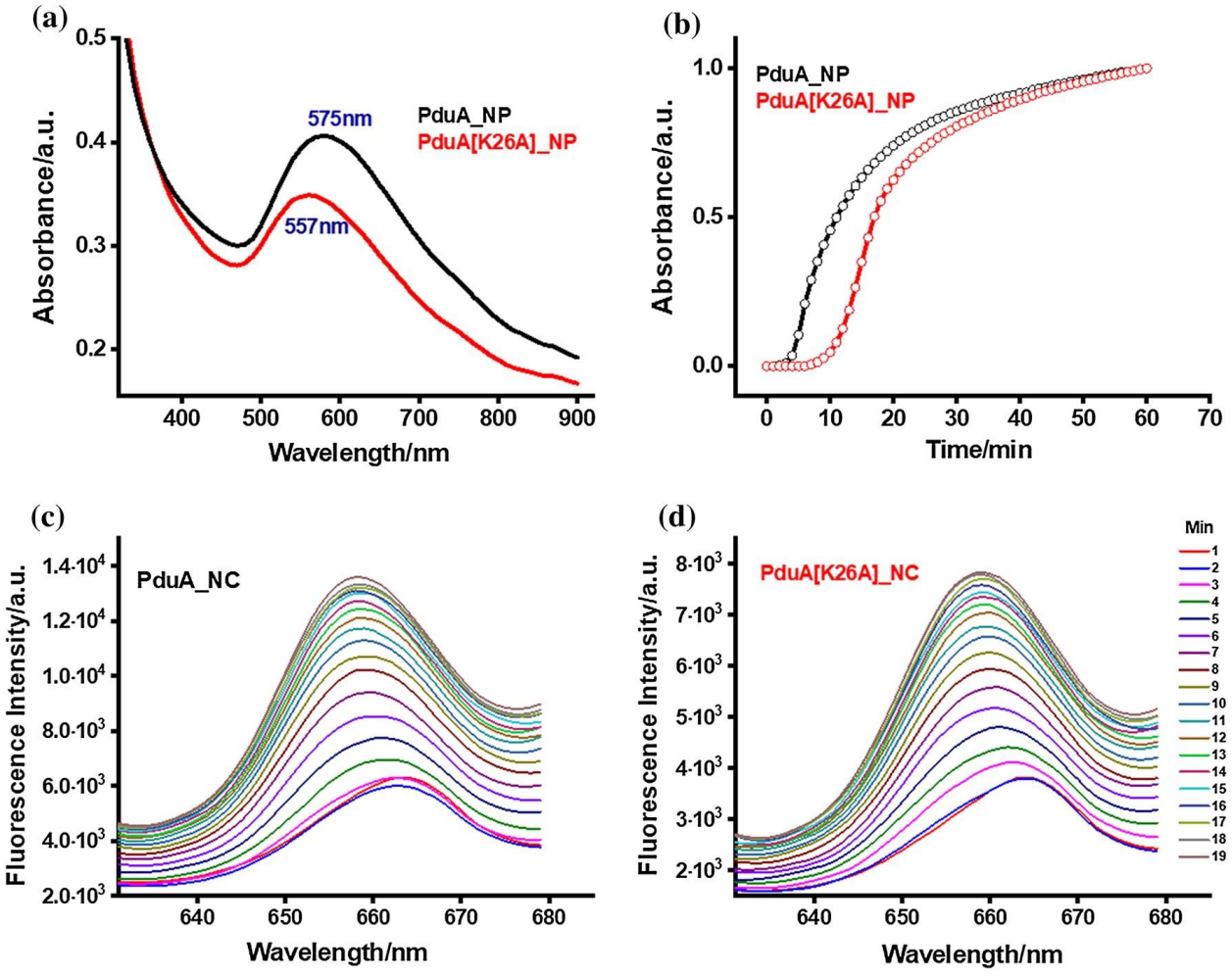
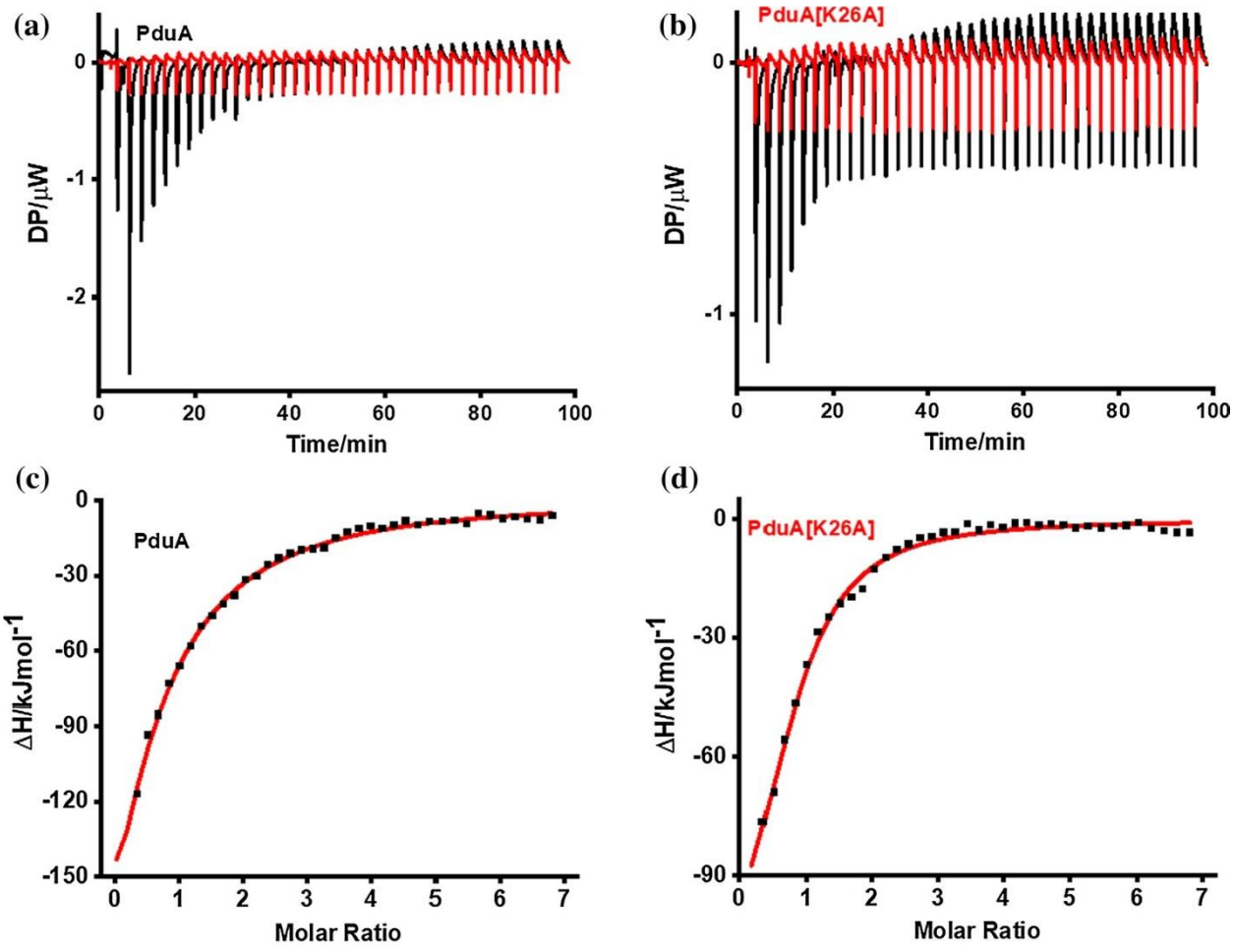


Figure 6.4 Synthesis of protein–gold nanoparticle conjugates. a) UV–Visible spectra showing a characteristic peak of fabricated gold nanoparticles in PduA [K26A]_NP (red) and a redshift in PduA_NP corresponding to large-size particles (black) b) Kinetics of reduction of gold nanoparticles on protein scaffolds PduA (black) and PduA [K26A] (red) with delayed initial absorbance. Fluorescence Spectra of gold nanoclusters (NCs) in the presence of c) PduA 2D and d) PduA[K26A] 3D protein scaffold showing nanocluster stage before nanoparticle formation (λ_{exc} 370 nm, λ_{em} 658 nm).

To probe the binding strength of interaction between gold ions and scaffold proteins, we performed Isothermal Titration Calorimetry (ITC). ITC titration is carried out at physiological pH 7.4 and room temperature 25 °C wherein 40 µl of 10 µM PduA and PduA[K26A] are injected in a constant volume (1 µl/injection) into 280 µl of 500 µM gold precursor solution in the reaction cell. Negative peaks in heat signal curves [Figure 6.5(a, b)] represent the exothermic process and integrated plots obtained from heat signal curves correspond to binding isotherms of gold to the protein surface. [Figure 6.5(c, d)] As seen from the binding isotherm plots, there is an interaction between the gold and scaffold proteins in a micro-molar range which signifies moderate binding without disturbing the protein scaffold native structure. The heat changes are analyzed after buffer subtraction to quantify the corresponding thermodynamic parameters of protein–gold ion interactions. Interaction of gold ions with PduA can fit at a mode of one site with N sites 2 suggesting gold ions bind to concave and convex faces with dissociation constant 45.0 µM. This complexation features favorable enthalpy change ($\Delta H = -250 \pm 17.7$) with $\Delta G = -24.8$ kJ/mol. In contrast, PduA [K26A] are fit at a mode of one site at binding affinity 6.6 µM with N site 1 which may be due to the inaccessible interacting sites resulting from 3D architecture [Figure. 6.5(e)]. This interaction is associated with lower enthalpy change $\Delta H = -105 \pm 4.4$ compared to the interaction with native PduA. This is also evident from our reduction experiments where PduA facilitates a faster reduction of gold nanoparticles.

6.4 Characterization of fabricated protein–nanoparticle conjugates

The native structure of the scaffold is an important parameter for the catalyst to function for efficient catalysis. Interaction of gold with proteins may affect their structure but some proteins can keep their native structures. PduA and PduA[K26A] self-assembling MCP shell protein scaffolds have a characteristic α-helix secondary structure as observed in circular dichroism (CD) spectra. The characteristic peak minima of α-helix are detected at wavelength 208 nm and 222 nm. Overlay of CD spectra of proteins alone and fabricated gold nanoparticles on the protein surface shows that the secondary structure is retained even after the fabrication of gold nanoparticles [Figure 6.6(a, b)]. There is a slight change in ellipticity values in both cases indicating the interaction of the nanoparticles with the protein surface. However, the overall conformation of



(e)

Parameters	PduA	PduA[K26A]
N	2.31±0.9	1.27 ± 0.1
Kd (M)	45.0e-6 ± 26.3e-6	6.64e-6 ± 1.41e-6
ΔH (kJ/mol)	-250 ± 17.7	-105 ± 4.40
-TΔS (kJ/K/mol)	225	75.7
ΔG (kJ/mol)	-24.8	-29.6

Figure 6.5 Isothermal titration calorimetry showing the interaction of $AuCl4^-$ with PduA and PduA [K26A] protein scaffolds. Raw heat change measured upon calorimetric titration of $AuCl4^-$ with protein scaffolds a) PduA b) PduA[K26A] showing saturating interaction. Buffer control is shown in red. Binding isotherm along with the fit (red) to raw data (black) using one site model for c) PduA d) PduA[K26A] to calculate the associated parameters after buffer subtraction.

the protein is not altered upon the fabrication of the nanoparticles. We observed that there is no detectable amount of free protein left in the solution post-fabrication of the nanoparticles [Figure 6.6 (a, b)]. The interaction of the gold ions with the protein surface as observed by ITC studies with micro-molar dissociation constants i.e., 45.0 µM for PduA and 6.64 µM for PduA[K26A] suggests moderate affinity which might not disrupt the protein structure and hence provides a stable protein–nanoparticle conjugate system [Figure 6.5].

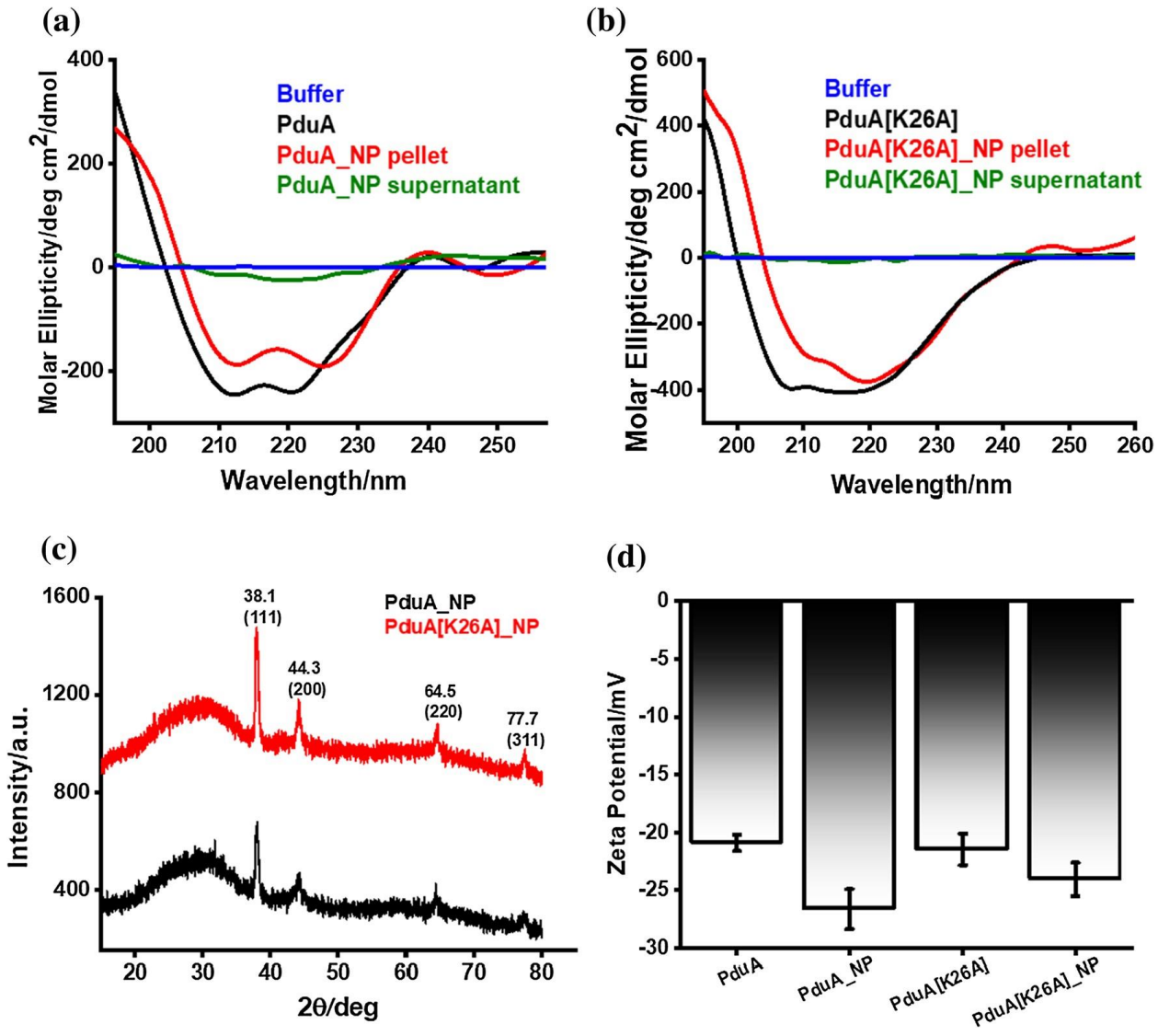


Figure 6.6 Characterization of protein–gold nanoparticle conjugates. CD spectra of proteins before and after fabrication of gold nanoparticles. The conjugated system was centrifuged and spectra recorded for supernatant and pellet re-dispersed a PduA b PduA[K26A]. Spectra show the retained secondary structure indicating the stability of the protein's structure post fabrication c XRD of crystalline gold nanoparticles PduA_NP (black) and PduA[K26A]_NP (red) d Zeta potential of protein scaffold alone and composite after fabrication of gold nanoparticles showing the stability of the colloidal system.

The hydrodynamic diameters of PduA and PduA[K26A] scaffold proteins without fabrication of gold as observed by dynamic light scattering (DLS) are in the range of 150–900 nm with (PDI < 0.5) for PduA and distribution for PduA[K26A] in the range 40–700 nm, respectively [Figure 6.7(a, b)]. This size distribution owes to the self-assembling tendency of proteins which has been observed earlier as well[174]. For protein–nanoparticle conjugates, it is difficult to measure the exact size of nanoparticles by DLS, so their size distribution is evaluated statistically by measuring the size diameter of approximately 150 nanoparticles from TEM images shown as a histogram plot [Figure 6.8(a, b)]. TEM images show that particles are spherical. The average particle size distribution of PduA[K26A]_NPs is 4.0 ± 0.07 nm with a narrow distribution from 2.5 nm to 5.5 nm whereas PduA_NPs range from 5.5 to 12.5 nm with an average particle size distribution of 8.4 ± 0.1 nm. The crystallinity of formed nanoparticles is determined by Powder X-ray diffraction. 2θ values of 38.1, 44.3, 64.5, and 77.7 indexed to (1 1 1), (2 0 0), (2 2 0), (3 1 1) reflection planes of face-centered cubic (fcc) structure of metallic gold [Figure 6.6(c)]. We also determined the zeta potential values of synthesized nanoparticles which is a measure of the stability of the colloidal solution. The zeta potential values near zero indicate no net charge and the least stable system. Here, the negative values of − 24 and − 26 for PduA and PduA[K26A] scaffold NP respectively, show the colloidal system is stable against aggregation. [Figure 6.6(d)].

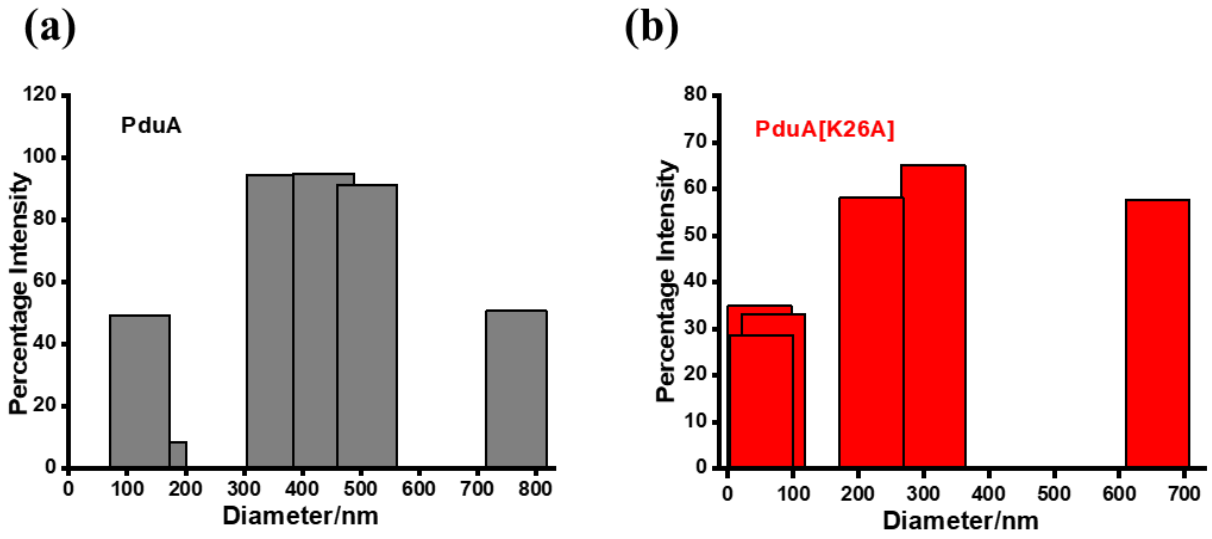


Figure 6.7 DLS showing size distribution for protein scaffolds a) PduA(black) b) PduA[K26A] (red). Protein scaffold alone in absence of gold fabrication showed size distribution in range from 150-900nm for PduA 2D sheets and distribution in PduA[K26A] in the range 40-700nm. This size distribution owes to the self-assembling tendency of proteins.

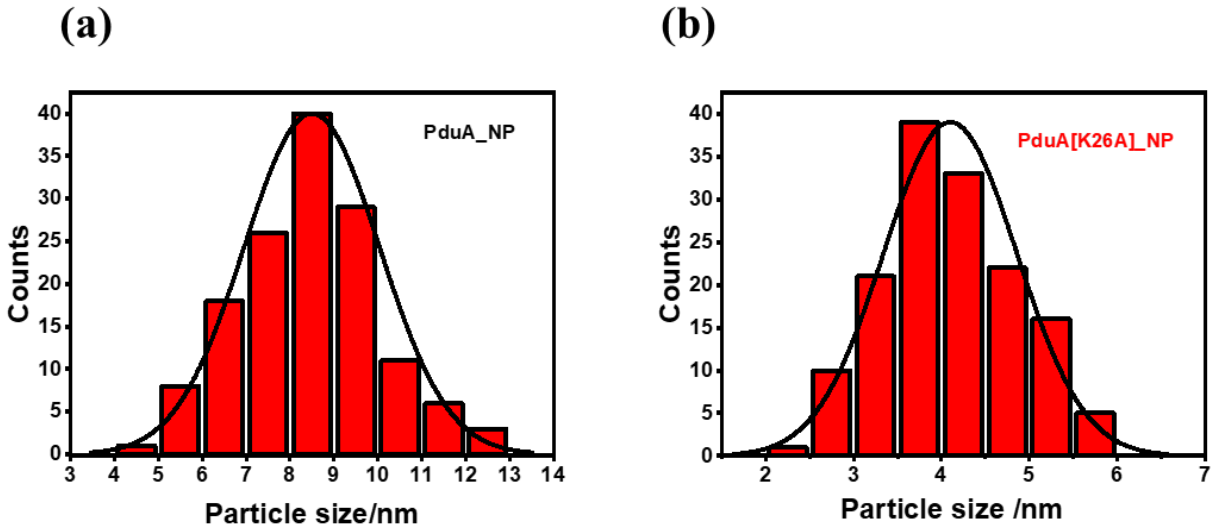


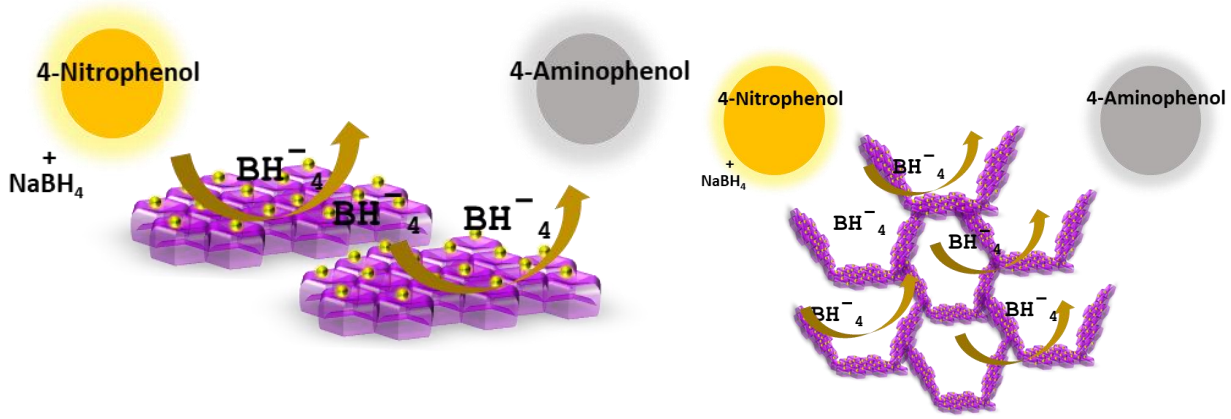
Figure 6.8 Histogram plots showing particle size distribution of protein-gold nanoparticles conjugates a) PduA NPs b) PduA[K26A] _NPs determined by their respective TEM images and fitted to Gaussian distribution. The average particle size distribution of PduA[K26A]

_NPs is 4.0±0.07nm with a narrow distribution from 2.5nm to 5.5nm whereas PduA NPs range from 5.5nm to 12.5nm with average particle size distribution of 8.4±0.1nm

6.5 Role of scaffold on the catalytic activity of protein– nanoparticle conjugates

6.5.1 Catalytic activity for reduction of 4-nitrophenol

We next, analyzed the catalytic efficiency of fabricated gold nanoparticles from 2D and 3D scaffolds using the catalytic reduction of 4-nitrophenol (4-NP) to 4-aminophenol (4-AP). This heterogeneous catalytic reaction is a well-known and generally selected model reaction for evaluating the catalytic activity of metal nanoparticles as catalysts[175–177]. Here, the reduction of 4-nitrophenol into 4-aminophenol by N aBH_4 is investigated. It has been suggested that gold nanoparticles catalyze this reaction by facilitating electron transfer from BH_4^- to 4-NP by adsorbing $NaBH_4$ on their surface. [Scheme 6.1] The characteristic absorption peak of 4-nitrophenol (4-NP) is observed at a wavelength of 317 nm. On the addition of $NaBH_4$, this peak is shifted to 400 nm due to the formation of 4-nitrophenolate anions. In the presence of catalytic AuNPs, 400 nm peak starts to decrease and a new peak at 300 nm simultaneously starts to appear which corresponds to the formation of 4-aminophenol.



Scheme 6.1 Mechanism of reduction of para-Nitrophenol by $NaBH_4$ to para-Aminophenol in the presence of the PduA_NPs and the PduA[K26A]_NPs.

In our experiment, we monitor the decrease in the absorption spectra at 400 nm and track the reaction for 20 min with a spectral reading from 250 to 600 nm wavelength [Figure 6.9(a, b)]. The decrease in the (4-NP) absorption peak and generation of (4-AP) absorption peak confirms the catalytic activity of as-synthesized gold nanoparticles. We found two-fold enhanced catalytic activity in the 3D scaffold. PduA[K26A] 3D scaffold nanoparticles deliver a higher catalytic rate with a rate constant (k) $3.0*10^{-1}$ min $^{-1}$ than the PduA 2D sheet scaffold (k) $1.7*10^{-1}$ min $^{-1}$ [Figure 6.10(a)], [Table 6.1]. This two-fold difference can be attributed to the spatial arrangement of small-size nanoparticles obtained from a 3D scaffold for the enhanced catalytic reaction.

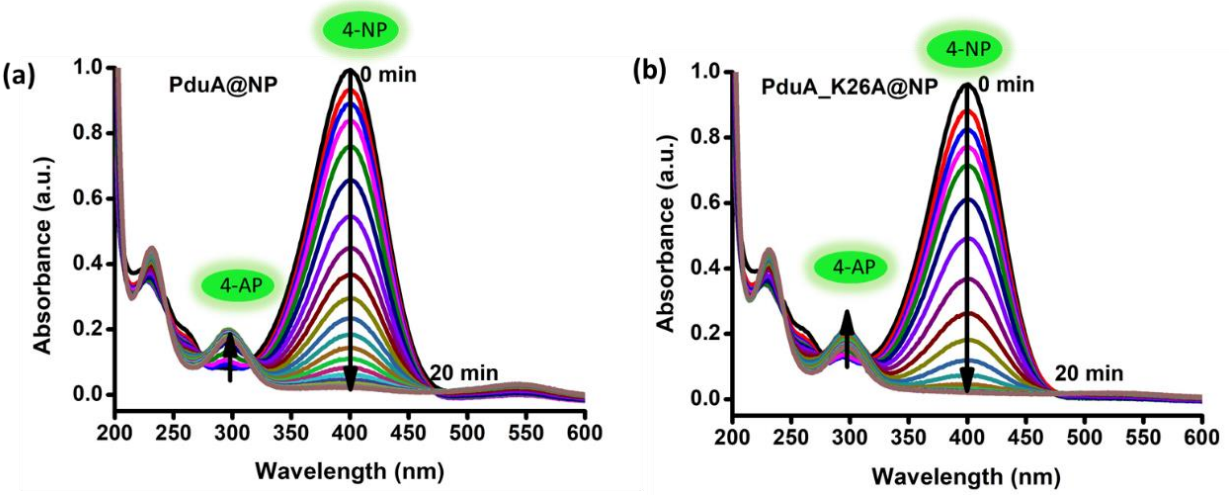


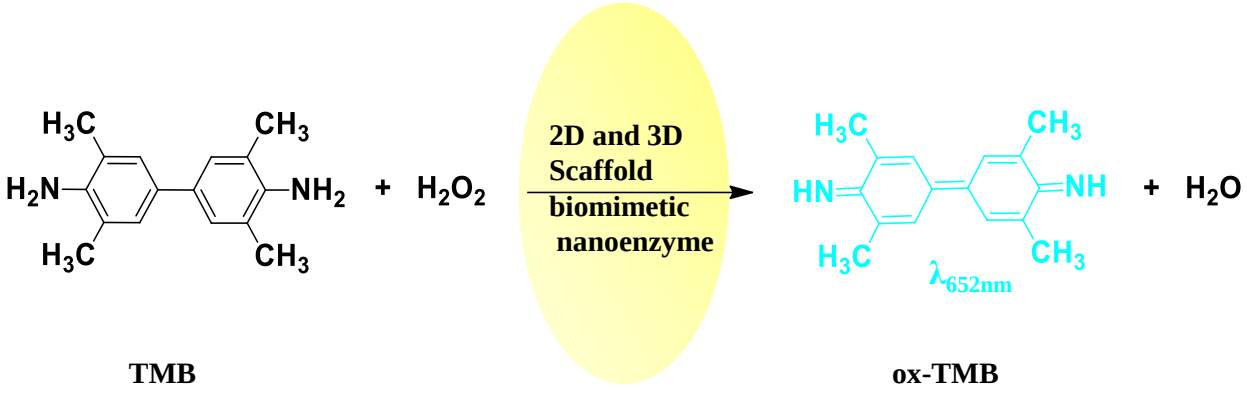
Figure 6.9 UV-Vis absorption spectra for conversion of 4-nitrophenol to 4-aminophenol. Reduction of 4-nitrophenolate ions by protein-gold nanoparticle conjugates a) PduA_NPs b) PduA[K26A] recorded at 1min interval for 20min.

Table 6.1 Catalytic rate constant for conversion of *p*-nitro-phenol to *p*-aminophenol with PduA_NPs and PduA[K26A]_NPs.

	PduA_NPs	PduA[K26A]_NPs
Rate constant (k)/min	1.7×10^{-1}	3.0×10^{-1}

6.5.2 The peroxidase mimicking the catalytic activity of gold nanoparticle–protein conjugates

There has been an increasing interest in artificial enzymes due to many shortcomings of natural enzymes, such as long-term stability and storage, easy denaturation, and highcost preparation and purification. Nanomaterials have proven to be attractive and highly efficient catalysts due to their large surface-to-volume ratio[137,138]. Gold nanoparticles have been demonstrated to behave like effective peroxidase enzyme mimetics (nanozymes)[178–181]. However, the catalytic efficiency is limited to acidic medium, high-temperature, surface coating to prevent aggregation of nanoparticles which is time-consuming and costly. Therefore, there is a need to explore and develop more stable nanoparticle systems with highly efficient catalytic activity.



Scheme 6.2 Oxidation of TMB substrate in the presence of 2D and 3D scaffolded biomimetic nanoenzymes to diamine product which absorbs at a wavelength of 652nm.

In this work, peroxidase mimicking catalytic activities of 2D scaffold and 3D scaffold gold nanoparticles have been explored and compared. To investigate the peroxidase-like activity of gold nanoparticles, we performed the steady-state kinetic assay at room temperature, under neutral

pH conditions. TMB (3,3′,5,5′-Tetramethylbenzidine) is used as a chromogenic substrate that has high sensitivity and easy colorimetric detection. TMB is oxidized by H_2O_2 and gives a diamine blue-colored product which absorbs at 652 nm wavelength[182] [Scheme 6.2]. We recorded the absorbance of the product formed in the presence of fabricated gold nanoparticles. However, we observed no color change in the absence of gold nanoparticles which served as a negative control. The kinetic parameters are obtained by varying the concentration of TMB and keeping the H_2O_2 concentration constant. Peroxidase catalysis by as-synthesized nanoparticles follows the Michaelis–Menten kinetics. The absorbance of the product formed at 652 nm is recorded and back-calculated to concentration using the Beer–Lambert Law and molar absorption coefficient of 39000 M^{-1} cm^{-1}. All reaction parameters are calculated using the Lineweaver–Burk equation to obtain V_{max} and K_m.

$$\frac{1}{V0} = \frac{K}{V\max m[s]} + \frac{1}{V\max 1} \qquad (1)$$

here, V_0 is the initial rate of reaction, V_{max} is the maximum conversion rate, [S] is the substrate concentration, and K_m is Michaelis constant which denotes the affinity of the enzyme.

The K_m and V_{max} values for PduA scaffold nanoparticles are 1.3 mM and $70.4*10^{-5}$ mM min^{-1} and for PduA[K26A] scaffold nanoparticles are 0.23 mM and $50.3*10^{-5}$ mM min^{-1} [Table 6.2]. Here, the small K_m (0.23 mM) for PduA[K26A] scaffold nanoparticles suggests a stronger affinity toward the substrate TMB than PduA (1.3 mM) scaffold nanoparticles.[Figure 6.10(b)]. This six-fold increased affinity would be a reason for the higher peroxidase activity of PduA[K26A] scaffold nanoparticles. The stronger affinity of PduA[K26A] is probably due to the high surface area of nanoparticles provided by this scaffold which in turn provides more active sites for TMB substrate to react as compared to its other analog. These results suggest that as-synthesized nanoparticles are potential enzyme mimetics for peroxidases. However, the effect of the 3D scaffold on the enhanced activity of synthesized nanoparticles is worth noticing with a six-fold increased affinity toward the substrate.

Table 6.2 Kinetic parameters for peroxidase mimicking activity of PduA_NPs and PduA[K26A]_NPs using TMB substrate.

	PduA_NPs	PduA[K26A]_NPs
V_{max} (mM/min)	70.3×10^{-5}	50.3×10^{-5}
K_m (mM)	1.3	0.23

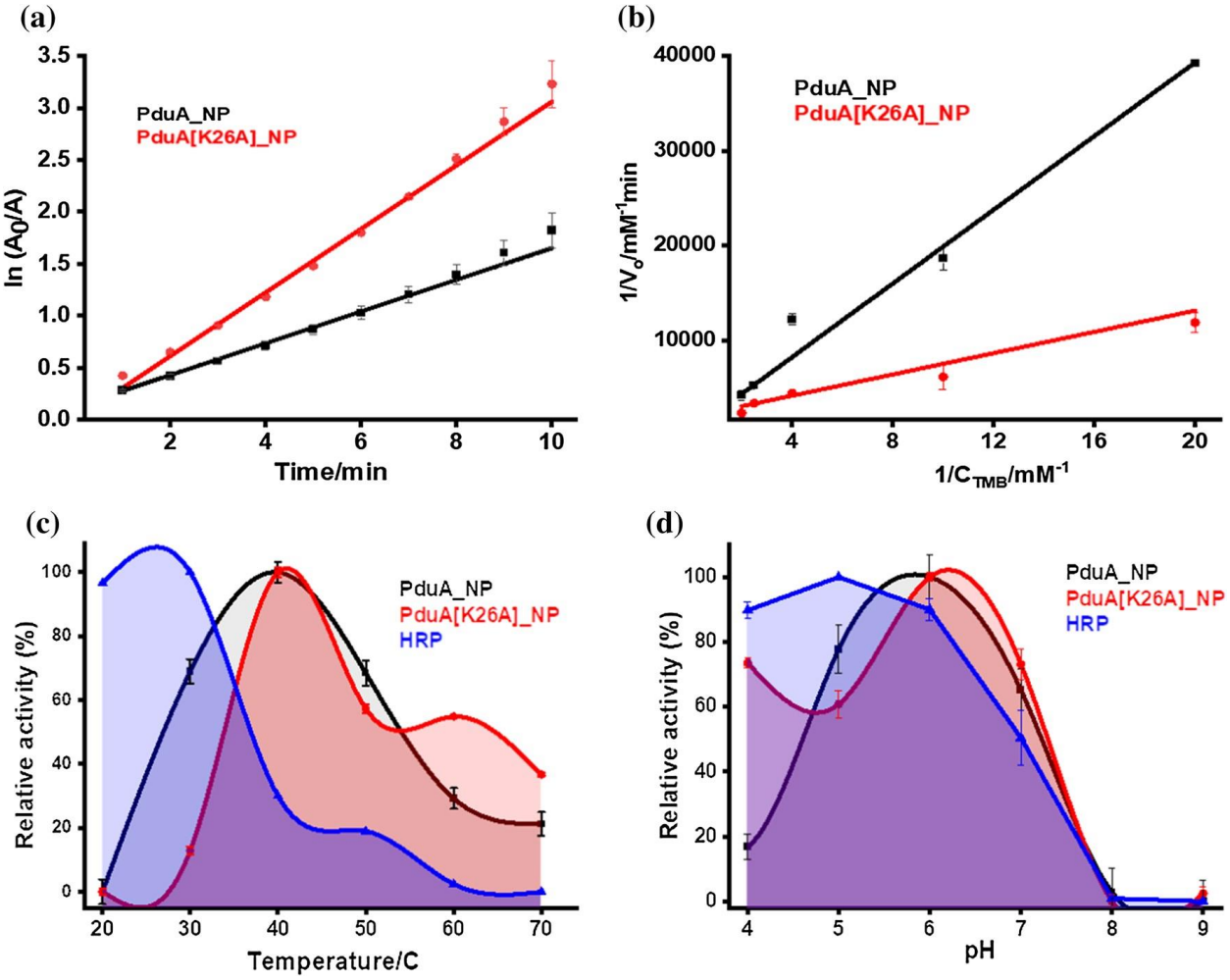


Figure 6.10 Protein–gold nanoparticle conjugates as efficient catalysis. a) Catalytic activity showing a two-fold increase in the presence of PduA[K26A] 3D scaffold (red) than PduA 2D sheet (black) for *para*nitro-phenol reduction. b) Lineweaver–Burk double reciprocal plot for

TMB enzyme kinetics with PduA (black) and PduA[K26A] (red) gold nanoparticles. The calculated V_{max} and K_m parameters show sixfold increase in affinity of PduA[K26A]_NP toward substrate due to small-size nanoparticles c) Temperature-dependant peroxidase activity of PduA_NP (black) and PduA[K26A]_NP (red) in comparison with native HRP enzyme (*blue*). d) pH-dependant peroxidase activity of PduA_NP (black) and PduA[K26A]_NP (red) in comparison with native HRP enzyme (blue). The highest enzyme activity is represented as 100% and relative activity values are calculated accordingly. Proteins due to their intrinsic thermal and pH stability retain more than 50% activity than native HRP enzyme at extreme temperature and pH conditions. Error bars shown represent the standard error of the mean from 3 experimental repeats.

6.5.3 Stability of biomimetic nanozymes

We also studied the robustness of as-synthesized nanoenzymes at a broader range of temperature and pH conditions. The peroxidase activity of nano-enzymes is performed and compared with the HRP (Horse Radish Peroxidase) enzyme by varying the temperature range from 20 to 80 °C and spectroscopically monitoring the change in absorbance at a wavelength of 652 nm. At every temperature, the nanoenzymes are incubated for 20 min and spectra are recorded to plot the relative product formation. There is a drastic decrease in catalytic activity after 40 °C in HRP control. However, nano-enzymes showed a wide range of catalytic activity, in the temperature range of 30–70 °C with the highest activity at 40 °C. The highest enzyme activity is represented as 100% and relative activity values are calculated accordingly. PduA [K26A]_NP retains 60% of maximum activity till 60 °C and drops to 40% at 70 °C, whereas PduA_ NP showed 20% less activity than PduA [K26A] in a similar temperature range [Figure 6.10(c)]. Similarly, we performed pH-dependent peroxidase activity in a pH range of 4–9 in glycine buffer and monitored spectroscopically. Both scaffold-derived nanoparticles are stable and retain more than 50% activity in a wide pH range of 4–7, whereas we found that PduA-derived nanoparticles showed less than 20% activity at pH 4. This observation correlates with the stability of the scaffold where proteins are found to be stable at these pH and temperature conditions[174]. On the other hand, HRP control lost its activity to 50% at pH 7 and completely loses at pH 8 [Figure 6.10(d)]. This suggests

that PduA and PduA[K26A] templated nanoparticles are more stable than native HRP at the wide distribution of pH and temperature. It is seen that these nanozymes have better stability and activity near physiological temperature and pH conditions. However, PduA[K26A] NP has better activity due to the large surface area of small-sized nanoparticles and its 3D architecture with the spatial distribution of nanoparticles is likely less prone to aggregation.

6.6 Conclusion

In the present study, we explored the intrinsic self-assembling nature of two shell proteins of MCP as a protein scaffold for the fabrication of gold nanoparticles. The hexameric shell protein, PduA forms stable 2D sheets and its mutant PduA[K26A] forms the 3D structure. These proteins have the same amino acid composition but are self-assembled into different hierarchical architectures due to only one amino acid change which makes them interesting and suitable to conduct these studies. PduA[K26A] 3D scaffold protein has a high surface area which resulted in the formation of more nucleation sites during the initial phase of the reduction process as reflected by the slow reduction rate and formation of small-size gold nanoparticles in comparison to the 2D scaffold. We observed that PduA scaffold interaction with gold ions is enthalpically favorable with two interacting sites compared to PduA[K26A] with one site as shown in ITC experiments. Further analysis of kinetic data of reduction of *para*-nitro-phenol catalytic activity of as-synthesized gold nanoparticles reveals a two-fold higher catalytic rate with 3D-scaffolded gold nanoparticles. This result is in accordance with the fact that PduA[K26A] nanoparticles are small in size and have a high spatial distribution which provides a high surface area for heterogeneous catalysis. Here we also report as-synthesized gold nanoparticles are potential nanoenzymes with peroxidase-like activity. In peroxidase activity, we observed that PduA[K26A] nanoparticles have a lower k_m value which is suggestive of high affinity toward the substrate. PduA[K26A]_NPs have six times higher affinity than PduA_NPs. Their comparison with standard HRP enzyme at wide pH and temperature range showed NP derived from protein scaffolds are more robust at extreme temperature and pH conditions. PduA[K26A] scaffold nanoparticles, however, have shown more than 50% activity at pH 4 and retain 60% activity at 60 °C and 70 °C whereas PduA nanoparticles retain only 20% and 40% at similar conditions respectively. Hence, we conclude that two different protein architect scaffolds built from the same building blocks result in different sizes and catalytic

efficiency of gold nanoparticles. From all these ensemble studies, it is quite evident that protein with a 3D scaffold serves as a better scaffold for fabrication and enhanced catalytic properties of gold nanoparticles. This is due to the spatial arrangement of gold nanoparticles on a high surface area of protein. This also highlights the fact that the choice of scaffold is an important measure to tune and tailor the properties of nanoparticles. Finally, the ease of genetically directing the scaffold structure and properties will allow further tunability of the shape, size, and overall structure characteristics of nanoparticles. Here, we explored these two self-assembling proteins in their different quaternary association as 2D and 3D scaffolds naturally to underlay the understanding of scaffold architecture on nanoparticle synthesis and properties. This paradigm will open room to develop new nanomaterial designs with desirable properties and answer questions in catalysis.

CHAPTER – 7

Summary and conclusion

Bacterial microcompartments (BMCs) are discovered as a unique nanomachinery fabricated from the self-assembly of proteins into a polyhedral architecture. The structure and the function of BMCs are both fascinating which has gathered a lot of attention towards it. Due to its all-protein nature, it is easily amendable and has entered the world of synthetic biology where protein-based nanoreactors and synthetic metabolic hubs can be designed. Studying and understanding microcompartments for such applications is essential. The complexity of the structure and function makes it difficult to solve its puzzle completely. Many studies during the past years have unfolded many curious aspects related to microcompartment structure and function. In this thesis, I have studied the sequence and structural diversity in the shell protein domains of the 1,2 propanediol bacterial microcompartment as a model system. Understanding the complex organization and architecture of BMC in this line will aid in understanding the evolutionary significance of diverse sequences and structures of shell proteins. Further, it will open a new paradigm for BMC proteins and can be explored for biomaterial and biotechnological applications.

To carry out the research work for this thesis I have explored molecular biology, biochemical and biophysical techniques. To decipher the structural complexity of shell proteins I choose two major shell proteins present in the Pdu operon (1,2 Pdu MCP) of *Salmonella enterica serovar typhimurium* which has different oligomerization states. These two shell proteins are present at the first two positions in an operon. PduA is present at the first position which oligomerizes as a hexamer from a single-domain monomer. Six such monomers self-assemble to form a hexamer. PduB' is another protein that occupies the second position in an operon and oligomerizes as a pseudohexamer from a dimer domain. Three such domains self-assemble to form a pseudohexamer (dimer-trimer). But interestingly, both proteins share similar hexagonal symmetry which is responsible for forming a 2D molecular sheet and hence the facets of microcompartment polyhedral structure. The shell proteins belong to the conserved family of bacterial microcompartment domain proteins (Pfam00936) and the variations at the core of the BMC domain have resulted in complicated fold variants that offer distinct properties and functions. The divergence in amino acid sequences, modular rearrangements in the structural domains, and circular permutations have increased the complexity of microcompartment structure. In ***Chapter 3***, I have developed synthetic dimer domain protein variants to understand the potential role of domain dimerization in MCP shell proteins. It is also related to the question of sequence and

structure variations. PduB' and PduT are two native dimer proteins present in MCP which have two fused circularly permuted domains and directly fused two domains of canonical domain fold respectively. Here I have used a native sequence of PduA single-domain protein and generated a dimer of PduA and circularly permuted dimer which is structurally similar to PduT and PduB' respectively. On characterization of proteins using circular dichroism and fluorescence spectroscopy, it was found proteins are correctly folded. Using dynamic light scattering and TEM micrographs, synthetic protein variants have shown remarkably different higher-order structures. PduA single domains when fused as PduA dimer, form tubes in contrast to PduA [WT] which forms 2D sheets. On circularly permuting and generating a dimer, it formed long thin sheets unlike PduB'[WT] which formed small sheets. In an approach to get insights into the important role of sequence in governing biophysical properties. Temperature-dependent studies showed the flexible nature of native dimer domain proteins where secondary structure elements uncoil at above 80^0C and regain the structure post-melt from circular dichroism spectroscopy. In contrast, synthetic dimer proteins don't show the uncoiling of secondary structure elements even at 95^0C. This characteristic is similar to the parent protein PduA. Hence, protein sequence governs the biophysical properties of BMC domain proteins and consecutively their functions as well.

In ***Chapter 4*** I have developed single-domain proteins as dimer-domain proteins have shown remarkable relatedness of shell proteins by sequence and structure. Here, the sequence from the native dimer domain protein was explored to generate a single-domain protein. PduB'[WT] was dissected into individual domains i.e., domain1 and domain2, and the second domain sequence was also explored to generate a PduA[WT] like single domain protein. Structural alignment of both domains (domain2 and PduA[WT]) showed circular permutation and accordingly, PduB' second domain was modified which has a similar topology to PduA[WT]. Protein variants were found to be correctly folded as assayed by circular dichroism and fluorescence experiments. DLS and TEM, evidently proved that they retained the self-assembling property but except PduB' domain1. Domain 2 has been shown to self-assemble to form long sheets whereas domain 2 with circular permutation (PduA[WT] like topology) could form small size sheets. Temperature-dependent results showed that the single-domain protein variants are thermostable and the secondary structure does not unfold to 95°C. Thus individual domains don't show a flexible nature as observed in native PduB'[WT]. This study suggests the presence of both domains together

governs the overall flexible structure and thus, provides resilience at a wide range of external stimuli.

In ***Chapter 5*** of this thesis, I have developed *in vivo* mutant strains where I have substituted the PduA [WT] protein with a structurally similar single domain protein obtained from PduB' domain2 after circular permutation. Also, PduA[WT] and PduBB'[WT] were swapped to develop a new mutant strain which has now PduBB' [WT] gene at the first position and PduA[WT] at the second position. For developing these *in vivo* strains I have explored and standardized a new powerful tool, CRISPR Cas9 genome editing. It was found that the shell protein replacements have a substantial effect on the growth of the bacteria determined by the growth curve of wild-type and modified strains in the presence of an inducer, 0.6% 1,2 propanediol (1,2 PD) supplemented with 150 nM CN-B_{12}. A longer lag phase and delayed growth in the mutant strain where PduA[WT] was substituted with its structurally similar variant indicated the toxicity produced from an intermediate, propanaldehyde. This could be because of the reason that the substituted variant could only form small sheets unlike PduA necessary for proper MCP formation. The other mutant where PduA: PduBB' was swapped also showed a long lag phase and delayed growth where later OD reaches similar to wild type but after 40h. However, despite many attempts, we could not purify MCP from both mutants. More insights can be delivered from more experiments and critical analysis in addition to these preliminary studies present in this thesis.

In ***Chapter 6***, I explored PduA shell protein for the development of hybrid nanocatalysts. The most stable and well-studied shell protein PduA and its mutant PduA[K26A] exhibit two different architectures. These naturally self-assembling proteins in 2D and 3D architectures reduce gold salt to gold nanoparticles and result in the different spatial organization of gold nanoparticles. TEM images and histogram plots showed different sizes of nanoparticles. The average particle size distribution of PduA[K26A]_NPs is 4.0 ± 0.07 nm with a narrow distribution from 2.5 nm to 5.5 nm whereas PduA_NPs range from 5.5 to 12.5 nm with an average particle size distribution of 8.4 ± 0.1 nm. Further protein nanoparticle conjugates when explored for catalytic efficiencies it is observed a two-fold increase in catalytic activity for the reduction of 4-nitrophenol. They have been shown to have remarkable peroxidase-mimicking catalytic activities. This approach demonstrates different architect scaffolds for the fabrication of gold nanoparticles can be explored

for tuning their properties and catalysis. It opens new paradigms toward the design of new nanomaterials and material science applications.

The key findings of the research work conducted in this thesis can be summarized in the following points. Native dimer proteins PduB'[WT] and PduT[WT] are dynamic structures that provide the BMC with a flexible structure. Single-domain proteins forming hexamers are more rigid and compact in structure in comparison to dimer-domain proteins which form pseudohexamer oligomerization. The single domains of PduB'[WT] protein work in a synergistic manner when present together. Dissected domains were found to be more compact whereas together they exhibit a dynamic structure. When synthetic proteins were accessed for the role of sequence and structure, dimers from the PduA[WT] sequence, irrespective of the topology attained, the properties belong to the original core of the amino acid sequence. These protein fold variants exhibited the evolutionary significant role of structural topology and their exquisite sequence specificities. The occurrences of various sequences in specific single and dimer domains provide BMC with global rigidity and flexibility to function in a better-coordinated manner. Lastly, the potential application of shell proteins has been shown, that were explored as scaffolds and development of nanocatalysts.

Bibliography

Bibliography

(1) Kerfeld, C. A.; Sawaya, M. R.; Tanaka, S.; Nguyen, C. V.; Phillips, M.; Beeby, M.; Yeates, T. O. Protein Structures Forming the Shell of Primitive Bacterial Organelles. *Science* **2005**, *309* (5736), 936–938. https://doi.org/10.1126/science.1113397.

(2) Axen, S. D.; Erbilgin, O.; Kerfeld, C. A. A Taxonomy of Bacterial Microcompartment Loci Constructed by a Novel Scoring Method. *PLoS Comput Biol* **2014**, *10* (10), e1003898. https://doi.org/10.1371/journal.pcbi.1003898.

(3) Chowdhury, C.; Sinha, S.; Chun, S.; Yeates, T. O.; Bobik, T. A. Diverse Bacterial Microcompartment Organelles. *Microbiol Mol Biol Rev* **2014**, *78* (3), 438–468. https://doi.org/10.1128/MMBR.00009-14.

(4) Shively, J. M.; Ball, F. L.; Kline, B. W. Electron Microscopy of the Carboxysomes (Polyhedral Bodies) of Thiobacillus Neapolitanus. *J Bacteriol* **1973**, *116* (3), 1405–1411.

(5) Kerfeld, C. A.; Melnicki, M. R. Assembly, Function and Evolution of Cyanobacterial Carboxysomes. *Curr Opin Plant Biol* **2016**, *31*, 66–75. https://doi.org/10.1016/j.pbi.2016.03.009.

(6) Kerfeld, C. A.; Erbilgin, O. Bacterial Microcompartments and the Modular Construction of Microbial Metabolism. *Trends Microbiol* **2015**, *23* (1), 22–34. https://doi.org/10.1016/j.tim.2014.10.003.

(7) *Ethanolamine Utilization in Bacteria | mBio*. https://journals.asm.org/doi/full/10.1128/mbio.00066-18 (accessed 2023-08-17).

(8) *Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica - PubMed*. https://pubmed.ncbi.nlm.nih.gov/16585748/ (accessed 2023-08-17).

(9) Rondon, M. R.; Escalante-Semerena, J. C. The Poc Locus Is Required for 1,2-Propanediol-Dependent Transcription of the Cobalamin Biosynthetic (Cob) and Propanediol Utilization (Pdu) Genes of Salmonella Typhimurium. *J Bacteriol* **1992**, *174* (7), 2267–2272. https://doi.org/10.1128/jb.174.7.2267-2272.1992.

(10) Cannon, G. C.; Bradburne, C. E.; Aldrich, H. C.; Baker, S. H.; Heinhorst, S.; Shively, J. M. Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra. *Appl Environ Microbiol* **2001**, *67* (12), 5351–5361. https://doi.org/10.1128/AEM.67.12.5351-5361.2001.

(11) Sutter, M.; Melnicki, M. R.; Schulz, F.; Woyke, T.; Kerfeld, C. A. A Catalog of the Diversity and Ubiquity of Bacterial Microcompartments. *Nat Commun* **2021**, *12* (1), 3809. https://doi.org/10.1038/s41467-021-24126-4.

(12) Craciun, S.; Balskus, E. P. Microbial Conversion of Choline to Trimethylamine Requires a Glycyl Radical Enzyme. *Proceedings of the National Academy of Sciences* **2012**, *109* (52), 21307–21312. https://doi.org/10.1073/pnas.1215689109.

(13) Herring, T. I.; Harris, T. N.; Chowdhury, C.; Mohanty, S. K.; Bobik, T. A. A Bacterial Microcompartment Is Used for Choline Fermentation by Escherichia Coli 536. *J Bacteriol* **2018**, *200* (10), e00764-17. https://doi.org/10.1128/JB.00764-17.

(14) Mallette, E.; Kimber, M. S. A Complete Structural Inventory of the Mycobacterial Microcompartment Shell Proteins Constrains Models of Global Architecture and Transport. *J Biol Chem* **2017**, *292* (4), 1197–1210. https://doi.org/10.1074/jbc.M116.754093.

(15) Petit, E.; LaTouf, W. G.; Coppi, M. V.; Warnick, T. A.; Currie, D.; Romashko, I.; Deshpande, S.; Haas, K.; Alvelo-Maurosa, J. G.; Wardman, C.; Schnell, D. J.; Leschine, S. B.; Blanchard, J. L. Involvement of a Bacterial Microcompartment in the Metabolism of Fucose and Rhamnose by Clostridium Phytofermentans. *PLOS ONE* **2013**, *8* (1), e54337. https://doi.org/10.1371/journal.pone.0054337.

(16) Zarzycki, J.; Sutter, M.; Cortina, N. S.; Erb, T. J.; Kerfeld, C. A. In Vitro Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment. *Sci Rep* **2017**, *7*, 42757. https://doi.org/10.1038/srep42757.

(17) Ferlez, B.; Sutter, M.; Kerfeld, C. A. Glycyl Radical Enzyme-Associated Microcompartments: Redox-Replete Bacterial Organelles. *mBio* **2019**, *10* (1), 10.1128/mbio.02327-18. https://doi.org/10.1128/mbio.02327-18.

(18) Lundin, A. P.; Stewart, K. L.; Stewart, A. M.; Herring, T. I.; Chowdhury, C.; Bobik, T. A. Genetic Characterization of a Glycyl Radical Microcompartment Used for 1,2-Propanediol Fermentation by Uropathogenic Escherichia Coli CFT073. *Journal of Bacteriology* **2020**, *202* (9), 10.1128/jb.00017-20. https://doi.org/10.1128/jb.00017-20.

(19) Cheng, S.; Liu, Y.; Crowley, C. S.; Yeates, T. O.; Bobik, T. A. Bacterial Microcompartments: Their Properties and Paradoxes. *Bioessays* **2008**, *30* (11–12), 1084–1095. https://doi.org/10.1002/bies.20830.

(20) Yeates, T. O.; Crowley, C. S.; Tanaka, S. Bacterial Microcompartment Organelles: Protein Shell Structure and Evolution. *Annu Rev Biophys* **2010**, *39*, 185–205. https://doi.org/10.1146/annurev.biophys.093008.131418.

(21) Drews, G.; Niklowitz, W. Beiträge zur Cytologie der Blaualgen. *Archiv. Mikrobiol.* **1956**, *24* (2), 147–162. https://doi.org/10.1007/BF00408629.

(22) Gantt, E.; Conti, S. F. Ultrastructure of Blue-Green Algae. *J Bacteriol* **1969**, *97* (3), 1486–1493. https://doi.org/10.1128/jb.97.3.1486-1493.1969.

(23) *Functional Organelles in Prokaryotes: Polyhedral Inclusions (Carboxysomes) of Thiobacillus neapolitanus | Science*. https://www.science.org/doi/10.1126/science.182.4112.584 (accessed 2023-07-01).

(24) Shively, J. M.; Ball, F.; Brown, D. H.; Saunders, R. E. Functional Organelles in Prokaryotes: Polyhedral Inclusions (Carboxysomes) of Thiobacillus Neapolitanus. *Science* **1973**, *182* (4112), 584–586. https://doi.org/10.1126/science.182.4112.584.

(25) Badger, M. R.; Price, G. D. The CO2concentrating Mechanism in Cyanobactiria and Microalgae. *Physiologia Plantarum* **1992**, *84* (4), 606–615. https://doi.org/10.1111/j.1399-3054.1992.tb04711.x.

(26) Kinney, J. N.; Axen, S. D.; Kerfeld, C. A. Comparative Analysis of Carboxysome Shell Proteins. *Photosynth Res* **2011**, *109* (1), 21–32. https://doi.org/10.1007/s11120-011-9624-6.

(27) Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C. Bacterial Microcompartments. *Annual Review of Microbiology* **2010**, *64* (1), 391–408. https://doi.org/10.1146/annurev.micro.112408.134211.

(28) Cannon, G. C.; Heinhorst, S.; Bradburne, C. E.; Shively, J. M. Carboxysome Genomics: A Status Report. *Functional Plant Biol.* **2002**, *29* (3), 175–182. https://doi.org/10.1071/pp01200.

(29) Badger, M. R.; Hanson, D.; Price, G. D. Evolution and Diversity of CO2 Concentrating Mechanisms in Cyanobacteria. *Funct Plant Biol* **2002**, *29* (3), 161–173. https://doi.org/10.1071/PP01213.

(30) Li, Y.; Kennedy, N. W.; Li, S.; Mills, C. E.; Tullman-Ercek, D.; Olvera de la Cruz, M. Computational and Experimental Approaches to Controlling Bacterial Microcompartment Assembly. *ACS Cent Sci* **2021**, *7* (4), 658–670. https://doi.org/10.1021/acscentsci.0c01699.

(31) Sampson, E. M.; Bobik, T. A. Microcompartments for B12-Dependent 1,2-Propanediol Degradation Provide Protection from DNA and Cellular Damage by a Reactive Metabolic Intermediate. *J Bacteriol* **2008**, *190* (8), 2966–2971. https://doi.org/10.1128/JB.01925-07.

(32) Yeates, T. O.; Thompson, M. C.; Bobik, T. A. The Protein Shells of Bacterial Microcompartment Organelles. *Current Opinion in Structural Biology* **2011**, *21* (2), 223–231. https://doi.org/10.1016/j.sbi.2011.01.006.

(33) Kerfeld, C. A.; Aussignargues, C.; Zarzycki, J.; Cai, F.; Sutter, M. Bacterial Microcompartments. *Nat Rev Microbiol* **2018**, *16* (5), 277–290. https://doi.org/10.1038/nrmicro.2018.10.

(34) Ochoa, J. M.; Yeates, T. O. Recent Structural Insights into Bacterial Microcompartment Shells. *Current Opinion in Microbiology* **2021**, *62*, 51–60. https://doi.org/10.1016/j.mib.2021.04.007.

(35) Yeates, T. O.; Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C.; Shively, J. M. Protein-Based Organelles in Bacteria: Carboxysomes and Related Microcompartments. *Nat Rev Microbiol* **2008**, *6* (9), 681–691. https://doi.org/10.1038/nrmicro1913.

(36) Tanaka, S.; Kerfeld, C. A.; Sawaya, M. R.; Cai, F.; Heinhorst, S.; Cannon, G. C.; Yeates, T. O. Atomic-Level Models of the Bacterial Carboxysome Shell. *Science* **2008**, *319* (5866), 1083–1086. https://doi.org/10.1126/science.1151458.

(37) Finn, R. D.; Tate, J.; Mistry, J.; Coggill, P. C.; Sammut, S. J.; Hotz, H.-R.; Ceric, G.; Forslund, K.; Eddy, S. R.; Sonnhammer, E. L. L.; Bateman, A. The Pfam Protein Families Database. *Nucleic Acids Res* **2008**, *36* (Database issue), D281-288. https://doi.org/10.1093/nar/gkm960.

(38) Krupovic, M.; Koonin, E. V. Cellular Origin of the Viral Capsid-like Bacterial Microcompartments. *Biology Direct* **2017**, *12* (1), 25. https://doi.org/10.1186/s13062-017-0197-y.

(39) Bobik, T. A.; Havemann, G. D.; Busch, R. J.; Williams, D. S.; Aldrich, H. C. The Propanediol Utilization (Pdu) Operon of Salmonella Enterica Serovar Typhimurium LT2 Includes Genes Necessary for Formation of Polyhedral Organelles Involved in Coenzyme B(12)-Dependent 1, 2-Propanediol Degradation. *J Bacteriol* **1999**, *181* (19), 5967–5975. https://doi.org/10.1128/JB.181.19.5967-5975.1999.

(40) Fan, C.; Cheng, S.; Sinha, S.; Bobik, T. A. Interactions between the Termini of Lumen Enzymes and Shell Proteins Mediate Enzyme Encapsulation into Bacterial Microcompartments. *Proc Natl Acad Sci U S A* **2012**, *109* (37), 14995–15000. https://doi.org/10.1073/pnas.1207516109.

(41) Crowley, C. S.; Sawaya, M. R.; Bobik, T. A.; Yeates, T. O. Structure of the PduU Shell Protein from the Pdu Microcompartment of Salmonella. *Structure* **2008**, *16* (9), 1324–1332. https://doi.org/10.1016/j.str.2008.05.013.

(42) Crowley, C. S.; Cascio, D.; Sawaya, M. R.; Kopstein, J. S.; Bobik, T. A.; Yeates, T. O. Structural Insight into the Mechanisms of Transport across the Salmonella Enterica Pdu Microcompartment Shell*. *Journal of Biological Chemistry* **2010**, *285* (48), 37838–37846. https://doi.org/10.1074/jbc.M110.160580.

(43) Tsai, Y.; Sawaya, M. R.; Cannon, G. C.; Cai, F.; Williams, E. B.; Heinhorst, S.; Kerfeld, C. A.; Yeates, T. O. Structural Analysis of CsoS1A and the Protein Shell of the Halothiobacillus Neapolitanus Carboxysome. *PLOS Biology* **2007**, *5* (6), e144. https://doi.org/10.1371/journal.pbio.0050144.

(44) Sinha, S.; Cheng, S.; Fan, C.; Bobik, T. A. The PduM Protein Is a Structural Component of the Microcompartments Involved in Coenzyme B12-Dependent 1,2-Propanediol Degradation by Salmonella Enterica. *Journal of Bacteriology* **2012**, *194* (8), 1912–1918. https://doi.org/10.1128/JB.06529-11.

(45) Sutter, M.; Faulkner, M.; Aussignargues, C.; Paasch, B. C.; Barrett, S.; Kerfeld, C. A.; Liu, L.-N. Visualization of Bacterial Microcompartment Facet Assembly Using High-Speed Atomic Force Microscopy. *Nano Lett* **2016**, *16* (3), 1590–1595. https://doi.org/10.1021/acs.nanolett.5b04259.

(46) Dryden, K. A.; Crowley, C. S.; Tanaka, S.; Yeates, T. O.; Yeager, M. Two-Dimensional Crystals of Carboxysome Shell Proteins Recapitulate the Hexagonal Packing of Three-Dimensional Crystals. *Protein Sci* **2009**, *18* (12), 2629–2635. https://doi.org/10.1002/pro.272.

(47) Sagermann, M.; Ohtaki, A.; Nikolakakis, K. Crystal Structure of the EutL Shell Protein of the Ethanolamine Ammonia Lyase Microcompartment. *Proc Natl Acad Sci U S A* **2009**, *106* (22), 8883–8887. https://doi.org/10.1073/pnas.0902324106.

(48) Ochoa, J. M.; Nguyen, V. N.; Nie, M.; Sawaya, M. R.; Bobik, T. A.; Yeates, T. O. Symmetry Breaking and Structural Polymorphism in a Bacterial Microcompartment Shell Protein for Choline Utilization. *Protein Science* **2020**, *29* (11), 2201–2212. https://doi.org/10.1002/pro.3941.

(49) Pang, A.; Warren, M. J.; Pickersgill, R. W. Structure of PduT, a Trimeric Bacterial Microcompartment Protein with a 4Fe-4S Cluster-Binding Site. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 2), 91–96. https://doi.org/10.1107/S0907444910050201.

(50) Heldt, D.; Frank, S.; Seyedarabi, A.; Ladikis, D.; Parsons, J. B.; Warren, M. J.; Pickersgill, R. W. Structure of a Trimeric Bacterial Microcompartment Shell Protein, EtuB, Associated with

Ethanol Utilization in Clostridium Kluyveri. *Biochem J* **2009**, *423* (2), 199–207. https://doi.org/10.1042/BJ20090780.

(51) Klein, M. G.; Zwart, P.; Bagby, S. C.; Cai, F.; Chisholm, S. W.; Heinhorst, S.; Cannon, G. C.; Kerfeld, C. A. Identification and Structural Analysis of a Novel Carboxysome Shell Protein with Implications for Metabolite Transport. *Journal of Molecular Biology* **2009**, *392* (2), 319–333. https://doi.org/10.1016/j.jmb.2009.03.056.

(52) Samborska, B.; Kimber, M. S. A Dodecameric CcmK2 Structure Suggests β-Carboxysomal Shell Facets Have a Double-Layered Organization. *Structure* **2012**, *20* (8), 1353–1362. https://doi.org/10.1016/j.str.2012.05.013.

(53) Cai, F.; Sutter, M.; Cameron, J. C.; Stanley, D. N.; Kinney, J. N.; Kerfeld, C. A. The Structure of CcmP, a Tandem Bacterial Microcompartment Domain Protein from the β-Carboxysome, Forms a Subcompartment within a Microcompartment. *J Biol Chem* **2013**, *288* (22), 16055–16063. https://doi.org/10.1074/jbc.M113.456897.

(54) Wheatley, N. M.; Gidaniyan, S. D.; Liu, Y.; Cascio, D.; Yeates, T. O. Bacterial Microcompartment Shells of Diverse Functional Types Possess Pentameric Vertex Proteins. *Protein Sci* **2013**, *22* (5), 660–665. https://doi.org/10.1002/pro.2246.

(55) Parsons, J. B.; Frank, S.; Bhella, D.; Liang, M.; Prentice, M. B.; Mulvihill, D. P.; Warren, M. J. Synthesis of Empty Bacterial Microcompartments, Directed Organelle Protein Incorporation, and Evidence of Filament-Associated Organelle Movement. *Mol Cell* **2010**, *38* (2), 305–315. https://doi.org/10.1016/j.molcel.2010.04.008.

(56) Mills, C. E.; Waltmann, C.; Archer, A. G.; Kennedy, N. W.; Abrahamson, C. H.; Jackson, A. D.; Roth, E. W.; Shirman, S.; Jewett, M. C.; Mangan, N. M.; Olvera de la Cruz, M.; Tullman-Ercek, D. Vertex Protein PduN Tunes Encapsulated Pathway Performance by Dictating Bacterial Metabolosome Morphology. *Nat Commun* **2022**, *13*, 3746. https://doi.org/10.1038/s41467-022-31279-3.

(57) Tanaka, S.; Sawaya, M. R.; Yeates, T. O. Structure and Mechanisms of a Protein-Based Organelle in Escherichia Coli. *Science* **2010**, *327* (5961), 81–84. https://doi.org/10.1126/science.1179513.

(58) Chen, P.; Andersson, D. I.; Roth, J. R. The Control Region of the Pdu/Cob Regulon in Salmonella Typhimurium. *J Bacteriol* **1994**, *176* (17), 5474–5482. https://doi.org/10.1128/jb.176.17.5474-5482.1994.

(59) Havemann, G. D.; Sampson, E. M.; Bobik, T. A. PduA Is a Shell Protein of Polyhedral Organelles Involved in Coenzyme B12-Dependent Degradation of 1,2-Propanediol in Salmonella Enterica Serovar Typhimurium LT2. *Journal of Bacteriology* **2002**, *184* (5), 1253–1261. https://doi.org/10.1128/JB.184.5.1253-1261.2002.

(60) AbdulRahman, F. The Distribution of Polyhedral Bacterial Microcompartments Suggests Frequent Horizontal Transfer and Operon Reassembly. In *Journal of Phylogenetics & Evolutionary Biology*; 2013; Vol. 01. https://doi.org/10.4172/2329-9002.1000118.

(61) Havemann, G. D.; Bobik, T. A. Protein Content of Polyhedral Organelles Involved in Coenzyme B12-Dependent Degradation of 1,2-Propanediol in Salmonella Enterica Serovar Typhimurium LT2. *J Bacteriol* **2003**, *185* (17), 5086–5095. https://doi.org/10.1128/JB.185.17.5086-5095.2003.

(62) Jorda, J.; Lopez, D.; Wheatley, N. M.; Yeates, T. O. Using Comparative Genomics to Uncover New Kinds of Protein-Based Metabolic Organelles in Bacteria. *Protein Sci* **2013**, *22* (2), 179–195. https://doi.org/10.1002/pro.2196.

(63) Jorda, J.; Liu, Y.; Bobik, T. A.; Yeates, T. O. Exploring Bacterial Organelle Interactomes: A Model of the Protein-Protein Interaction Network in the Pdu Microcompartment. *PLOS Computational Biology* **2015**, *11* (2), e1004067. https://doi.org/10.1371/journal.pcbi.1004067.

(64) Chowdhury, C.; Chun, S.; Sawaya, M. R.; Yeates, T. O.; Bobik, T. A. The Function of the PduJ Microcompartment Shell Protein Is Determined by the Genomic Position of Its Encoding Gene. *Mol Microbiol* **2016**, *101* (5), 770–783. https://doi.org/10.1111/mmi.13423.

(65) Cheng, S.; Sinha, S.; Fan, C.; Liu, Y.; Bobik, T. A. Genetic Analysis of the Protein Shell of the Microcompartments Involved in Coenzyme B12-Dependent 1,2-Propanediol Degradation by Salmonella. *J Bacteriol* **2011**, *193* (6), 1385–1392. https://doi.org/10.1128/JB.01473-10.

(66) Chowdhury, C.; Chun, S.; Pang, A.; Sawaya, M. R.; Sinha, S.; Yeates, T. O.; Bobik, T. A. Selective Molecular Transport through the Protein Shell of a Bacterial Microcompartment Organelle. *Proceedings of the National Academy of Sciences* **2015**, *112* (10), 2990–2995. https://doi.org/10.1073/pnas.1423672112.

(67) Park, J.; Chun, S.; Bobik, Thomas. A.; Houk, K. N.; Yeates, T. O. Molecular Dynamics Simulations of Selective Metabolite Transport across the Propanediol Bacterial Microcompartment Shell. *J Phys Chem B* **2017**, *121* (34), 8149–8154. https://doi.org/10.1021/acs.jpcb.7b07232.

(68) Slininger Lee, M. F.; Jakobson, C. M.; Tullman-Ercek, D. Evidence for Improved Encapsulated Pathway Behavior in a Bacterial Microcompartment through Shell Protein Engineering. *ACS Synth. Biol.* **2017**, *6* (10), 1880–1891. https://doi.org/10.1021/acssynbio.7b00042.

(69) Uddin, I.; Frank, S.; Warren, M. J.; Pickersgill, R. W. A Generic Self-Assembly Process in Microcompartments and Synthetic Protein Nanotubes. *Small* **2018**, *14* (19), 1704020. https://doi.org/10.1002/smll.201704020.

(70) Pang, A.; Frank, S.; Brown, I.; Warren, M. J.; Pickersgill, R. W. Structural Insights into Higher Order Assembly and Function of the Bacterial Microcompartment Protein PduA. *J Biol Chem* **2014**, *289* (32), 22377–22384. https://doi.org/10.1074/jbc.M114.569285.

(71) Parsons, J. B.; Dinesh, S. D.; Deery, E.; Leech, H. K.; Brindley, A. A.; Heldt, D.; Frank, S.; Smales, C. M.; Lünsdorf, H.; Rambach, A.; Gass, M. H.; Bleloch, A.; McClean, K. J.; Munro, A. W.; Rigby, S. E. J.; Warren, M. J.; Prentice, M. B. Biochemical and Structural Insights into Bacterial Organelle Form and Biogenesis. *J Biol Chem* **2008**, *283* (21), 14366–14375. https://doi.org/10.1074/jbc.M709214200.

(72) Kennedy, N. W.; Mills, C. E.; Abrahamson, C. H.; Archer, A. G.; Shirman, S.; Jewett, M. C.; Mangan, N. M.; Tullman-Ercek, D. Linking the Salmonella Enterica 1,2-Propanediol Utilization Bacterial Microcompartment Shell to the Enzymatic Core via the Shell Protein PduB. *J Bacteriol* **2022**, *204* (9), e0057621. https://doi.org/10.1128/jb.00576-21.

(73) Lehman, B. P.; Chowdhury, C.; Bobik, T. A. 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), e00785-16. https://doi.org/10.1128/JB.00785-16.

(74) Kumar, G.; Hazra, J. P.; Sinha, S. Disordered Regions Endow Structural Flexibility to Shell Proteins and Function towards Shell-Enzyme Interactions in 1,2-Propanediol Utilization Microcompartment. *J Biomol Struct Dyn* **2022**, 1–11. https://doi.org/10.1080/07391102.2022.2138552.

(75) Larsson, A. M.; Hasse, D.; Valegård, K.; Andersson, I. Crystal Structures of β-Carboxysome Shell Protein CcmP: Ligand Binding Correlates with the Closed or Open Central Pore. *Journal of Experimental Botany* **2017**, *68* (14), 3857. https://doi.org/10.1093/jxb/erx070.

(76) Pang, A.; Liang, M.; Prentice, M. B.; Pickersgill, R. W. Substrate Channels Revealed in the Trimeric Lactobacillus Reuteri Bacterial Microcompartment Shell Protein PduB. *Acta*

Crystallogr D Biol Crystallogr **2012**, *68* (Pt 12), 1642–1652. https://doi.org/10.1107/S0907444912039315.

(77) Yeates, T. O.; Jorda, J.; Bobik, T. A. The Shells of BMC-Type Microcompartment Organelles in Bacteria. *J Mol Microbiol Biotechnol* **2013**, *23* (4–5), 290–299. https://doi.org/10.1159/000351347.

(78) Frank, S.; Lawrence, A. D.; Prentice, M. B.; Warren, M. J. Bacterial Microcompartments Moving into a Synthetic Biological World. *Journal of Biotechnology* **2013**, *163* (2), 273–279. https://doi.org/10.1016/j.jbiotec.2012.09.002.

(79) Sinha, S.; Cheng, S.; Sung, Y. W.; McNamara, D. E.; Sawaya, M. R.; Yeates, T. O.; Bobik, T. A. Alanine Scanning Mutagenesis Identifies an Asparagine-Arginine-Lysine Triad Essential to Assembly of the Shell of the Pdu Microcompartment. *J Mol Biol* **2014**, *426* (12), 2328–2345. https://doi.org/10.1016/j.jmb.2014.04.012.

(80) Kaur, S.; Bari, N. K.; Sinha, S. Varying Protein Architectures in 3-Dimensions for Scaffolding and Modulating Properties of Catalytic Gold Nanoparticles. *Amino Acids* **2022**, *54* (3), 441–454. https://doi.org/10.1007/s00726-022-03127-7.

(81) Bari, N. K.; Hazra, J. P.; 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. *Biochim Biophys Acta Gen Subj* **2020**, *1864* (10), 129680. https://doi.org/10.1016/j.bbagen.2020.129680.

(82) Link, S.; El-Sayed, M. A. Optical Properties and Ultrafast Dynamics of Metallic Nanocrystals. *Annual Review of Physical Chemistry* **2003**, *54* (1), 331–366. https://doi.org/10.1146/annurev.physchem.54.011002.103759.

(83) Underwood, S.; Mulvaney, P. Effect of the Solution Refractive Index on the Color of Gold Colloids. *Langmuir* **1994**, *10* (10), 3427–3430. https://doi.org/10.1021/la00022a011.

(84) Zhao, P.; Feng, X.; Huang, D.; Yang, G.; Astruc, D. Basic Concepts and Recent Advances in Nitrophenol Reduction by Gold- and Other Transition Metal Nanoparticles. *Coordination Chemistry Reviews* **2015**, *C* (287), 114–136. https://doi.org/10.1016/j.ccr.2015.01.002.

(85) Aditya, T.; Pal, A.; Pal, T. Nitroarene Reduction: A Trusted Model Reaction to Test Nanoparticle Catalysts. *Chem. Commun.* **2015**, *51* (46), 9410–9431. https://doi.org/10.1039/C5CC01131K.

(86) Josephy, P. D.; Eling, T.; Mason, R. P. The Horseradish Peroxidase-Catalyzed Oxidation of 3,5,3',5'-Tetramethylbenzidine. Free Radical and Charge-Transfer Complex Intermediates. *Journal of Biological Chemistry* **1982**, *257* (7), 3669–3675. https://doi.org/10.1016/S0021-9258(18)34832-4.

(87) *Schrodinger, L. (2010) The PyMOL Molecular Graphics System, Version 1.3r1. - Google Search*. https://www.google.com/search?q=Schrodinger%2C+L.+(2010)+The+PyMOL+Molecular+Graphics+System%2C+Version+1.3r1.&rlz=1C1CHBD_enIN919IN919&oq=Schrodinger%2C+L.+(2010)+The+PyMOL+Molecular+Graphics+System%2C+Version+1.3r1.&aqs=chrome..69i57j69i61.673j0j4&sourceid=chrome&ie=UTF-8 (accessed 2023-04-26).

(88) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat Methods* **2009**, *6* (5), 343–345. https://doi.org/10.1038/nmeth.1318.

(89) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(90) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat Methods* **2022**, *19* (6), 679–682. https://doi.org/10.1038/s41592-022-01488-1.

(91) Kuriata, A.; Gierut, A. M.; Oleniecki, T.; Ciemny, M. P.; 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. https://doi.org/10.1093/nar/gky356.

(92) Vander Meersche, Y.; Cretin, G.; de Brevern, A. G.; Gelly, J.-C.; Galochkina, T. MEDUSA: Prediction of Protein Flexibility from Sequence. *Journal of Molecular Biology* **2021**, *433* (11), 166882. https://doi.org/10.1016/j.jmb.2021.166882.

(93) Zhang, H.; Cheng, Q.-X.; Liu, A.-M.; Zhao, G.-P.; Wang, J. A Novel and Efficient Method for Bacteria Genome Editing Employing Both CRISPR/Cas9 and an Antibiotic Resistance Cassette. *Front Microbiol* **2017**, *8*, 812. https://doi.org/10.3389/fmicb.2017.00812.

(94) Sutter, M.; Greber, B.; Aussignargues, C.; Kerfeld, C. A. Assembly Principles and Structure of a 6.5 MDa Bacterial Microcompartment Shell. *Science (New York, N.Y.)* **2017**, *356* (6344), 1293. https://doi.org/10.1126/science.aan3289.

(95) Tanaka, S.; Sawaya, M. R.; Yeates, T. O. Structure and Mechanisms of a Protein-Based Organelle in Escherichia Coli. *Science* **2010**, *327* (5961), 81–84. https://doi.org/10.1126/science.1179513.

(96) Ferlez, B. H.; Kirst, H.; Greber, B. J.; Nogales, E.; Sutter, M.; Kerfeld, C. A. Heterologous Assembly of Pleomorphic Bacterial Microcompartment Shell Architectures Spanning the Nano- to Microscale. *Advanced Materials n/a* (n/a), 2212065. https://doi.org/10.1002/adma.202212065.

(97) Yeates, T. O.; Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C.; Shively, J. M. Protein-Based Organelles in Bacteria: Carboxysomes and Related Microcompartments. *Nat Rev Microbiol* **2008**, *6* (9), 681–691. https://doi.org/10.1038/nrmicro1913.

(98) Pang, A.; Warren, M. J.; Pickersgill, R. W. Structure of PduT, a Trimeric Bacterial Microcompartment Protein with a 4Fe-4S Cluster-Binding Site. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 2), 91–96. https://doi.org/10.1107/S0907444910050201.

(99) Yeates, T. O.; Crowley, C. S.; Tanaka, S. Bacterial Microcompartment Organelles: Protein Shell Structure and Evolution. *Annu Rev Biophys* **2010**, *39*, 185–205. https://doi.org/10.1146/annurev.biophys.093008.131418.

(100) Bliven, S.; Prlić, A. Circular Permutation in Proteins. *PLOS Computational Biology* **2012**, *8* (3), e1002445. https://doi.org/10.1371/journal.pcbi.1002445.

(101) Cunningham, B. A.; Hemperly, J. J.; Hopp, T. P.; Edelman, G. M. Favin versus Concanavalin A: Circularly Permuted Amino Acid Sequences. *Proceedings of the National Academy of Sciences* **1979**, *76* (7), 3218–3222. https://doi.org/10.1073/pnas.76.7.3218.

(102) Weiner, J., 3rd; Bornberg-Bauer, E. Evolution of Circular Permutations in Multidomain Proteins. *Molecular Biology and Evolution* **2006**, *23* (4), 734–743. https://doi.org/10.1093/molbev/msj091.

(103) Whitehead, T. A.; Bergeron, L. M.; Clark, D. S. Tying up the Loose Ends: Circular Permutation Decreases the Proteolytic Susceptibility of Recombinant Proteins. *Protein Eng Des Sel* **2009**, *22* (10), 607–613. https://doi.org/10.1093/protein/gzp034.

(104) Ostermeier, M. Engineering Allosteric Protein Switches by Domain Insertion. *Protein Eng Des Sel* **2005**, *18* (8), 359–364. https://doi.org/10.1093/protein/gzi048.

(105) Yu, Y.; Lutz, S. Circular Permutation: A Different Way to Engineer Enzyme Structure and Function. *Trends Biotechnol* **2011**, *29* (1), 18–25. https://doi.org/10.1016/j.tibtech.2010.10.004.

(106) Luger, K.; Hommel, U.; Herold, M.; Hofsteenge, J.; Kirschner, K. Correct Folding of Circularly Permuted Variants of a Beta Alpha Barrel Enzyme in Vivo. *Science* **1989**, *243* (4888), 206–210. https://doi.org/10.1126/science.2643160.

(107) Qian, Z.; Lutz, S. Improving the Catalytic Activity of Candida Antarctica Lipase B by Circular Permutation. *J Am Chem Soc* **2005**, *127* (39), 13466–13467. https://doi.org/10.1021/ja053932h.

(108) Kostyuk, A. I.; Demidovich, A. D.; Kotova, D. A.; Belousov, V. V.; Bilan, D. S. Circularly Permuted Fluorescent Protein-Based Indicators: History, Principles, and Classification. *Int J Mol Sci* **2019**, *20* (17), 4200. https://doi.org/10.3390/ijms20174200.

(109) Bari, N. K.; Kumar, G.; Hazra, J. P.; Kaur, S.; Sinha, S. Functional Protein Shells Fabricated from the Self-Assembling Protein Sheets of Prokaryotic Organelles. *J. Mater. Chem. B* **2020**, *8* (3), 523–533. https://doi.org/10.1039/C9TB02224D.

(110) 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 Res Lett* **2019**, *14* (1), 54. https://doi.org/10.1186/s11671-019-2884-3.

(111) Kennedy, N. W.; Ikonomova, S. P.; Slininger Lee, M.; Raeder, H. W.; Tullman-Ercek, D. Self-Assembling Shell Proteins PduA and PduJ Have Essential and Redundant Roles in Bacterial Microcompartment Assembly. *J Mol Biol* **2021**, *433* (2), 166721. https://doi.org/10.1016/j.jmb.2020.11.020.

(112) Kumar, G.; Bari, N. K.; Hazra, J. P.; Sinha, S. A Major Shell Protein of 1,2-Propanediol Utilization Microcompartment Conserves the Activity of Its Signature Enzyme at Higher Temperatures. *ChemBioChem* **2022**, *23* (9), e202100694. https://doi.org/10.1002/cbic.202100694.

(113) Niesen, F. H.; Berglund, H.; Vedadi, M. The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability. *Nat Protoc* **2007**, *2* (9), 2212–2221. https://doi.org/10.1038/nprot.2007.321.

(114) Williams, S. G.; Lovell, S. C. The Effect of Sequence Evolution on Protein Structural Divergence. *Molecular Biology and Evolution* **2009**, *26* (5), 1055–1065. https://doi.org/10.1093/molbev/msp020.

(115) Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nat Rev Mol Cell Biol* **2019**, *20* (11), 681–697. https://doi.org/10.1038/s41580-019-0163-x.

(116) Bashton, M.; Chothia, C. The Generation of New Protein Functions by the Combination of Domains. *Structure* **2007**, *15* (1), 85–99. https://doi.org/10.1016/j.str.2006.11.009.

(117) Bornberg-Bauer, E.; Albà, M. M. Dynamics and Adaptive Benefits of Modular Protein Evolution. *Curr Opin Struct Biol* **2013**, *23* (3), 459–466. https://doi.org/10.1016/j.sbi.2013.02.012.

(118) Chothia, C.; Gough, J.; Vogel, C.; Teichmann, S. A. Evolution of the Protein Repertoire. *Science* **2003**, *300* (5626), 1701–1703. https://doi.org/10.1126/science.1085371.

(119) Wang, M.; Caetano-Anollés, G. The Evolutionary Mechanics of Domain Organization in Proteomes and the Rise of Modularity in the Protein World. *Structure* **2009**, *17* (1), 66–78. https://doi.org/10.1016/j.str.2008.11.008.

(120) Price, M. N.; Arkin, A. P.; Alm, E. J. The Life-Cycle of Operons. *PLOS Genetics* **2006**, *2* (6), e96. https://doi.org/10.1371/journal.pgen.0020096.

(121) Kovács, K.; Hurst, L. D.; Papp, B. Stochasticity in Protein Levels Drives Colinearity of Gene Order in Metabolic Operons of Escherichia Coli. *PLoS Biol* **2009**, *7* (5), e1000115. https://doi.org/10.1371/journal.pbio.1000115.

(122) Jorda, J.; Leibly, D. J.; Thompson, M. C.; Yeates, T. O. Structure of a Novel 13 Nm Dodecahedral Nanocage Assembled from a Redesigned Bacterial Microcompartment Shell Protein. *Chem. Commun.* **2016**, *52* (28), 5041–5044. https://doi.org/10.1039/C6CC00851H.

(123) Lee, M. J.; Mantell, J.; Brown, I. R.; Fletcher, J. M.; Verkade, P.; Pickersgill, R. W.; Woolfson, D. N.; Frank, S.; Warren, M. J. De Novo Targeting to the Cytoplasmic and Luminal Side of Bacterial Microcompartments. *Nat Commun* **2018**, *9* (1), 3413. https://doi.org/10.1038/s41467-018-05922-x.

(124) Ferlez, B.; Sutter, M.; Kerfeld, C. A. A Designed Bacterial Microcompartment Shell with Tunable Composition and Precision Cargo Loading. *Metab Eng* **2019**, *54*, 286–291. https://doi.org/10.1016/j.ymben.2019.04.011.

(125) Sinha, S.; Cheng, S.; Sung, Y. W.; Mcnamara, D. E.; Sawaya, M. R.; Yeates, T. O.; Bobik, T. A.; Carver, R. J. Alanine Scanning Mutagenesis Identifies an Asparagine- Arginine-Lysine Triad Essential to Assembly of the Shell of the Pdu Microcompartment The Coordinates of the Refined Models and the Merged Structure Factors Have Been Deposited in the Protein Data Bank. *J Mol Biol. J Mol Biol. June* **2014**, *12* (42612), 2328–2345. https://doi.org/10.1016/j.jmb.2014.04.012.

(126) Salata, O. V. Applications of Nanoparticles in Biology and Medicine. *Journal of Nanobiotechnology* **2004**, *2*, 1–6. https://doi.org/10.1186/1477-3155-2-3.

(127) Loizidou, M.; Seifalian, A. M. Nanotechnology and Its Applications in Surgery. *British Journal of Surgery* **2010**, *97* (4), 463–465. https://doi.org/10.1002/bjs.7074.

(128) Martis, E.; Badve, R.; Degwekar, M. Nanotechnology Based Devices and Applications in Medicine: An Overview. *Chronicles of Young Scientists* **2012**, *3* (1), 68. https://doi.org/10.4103/2229-5186.94320.

(129) Nie, L.; Liu, F.; Ma, P.; Xiao, X. Applications of Gold Nanoparticles in Optical Biosensors. *Journal of Biomedical Nanotechnology* **2014**, *10* (10), 2700–2721. https://doi.org/10.1166/jbn.2014.1987.

(130) Wen, W.; Song, Y.; Yan, X.; Zhu, C.; Du, D.; Wang, S.; Asiri, A. M.; Lin, Y. Recent Advances in Emerging 2D Nanomaterials for Biosensing and Bioimaging Applications. *Materials Today* **2018**, *21* (2), 164–177. https://doi.org/10.1016/j.mattod.2017.09.001.

(131) Chen, J.; Meng, H.; Tian, Y.; Yang, R.; Du, D.; Li, Z.; Qu, L.; Lin, Y. Recent Advances in Functionalized MnO 2 Nanosheets for Biosensing and Biomedicine Applications. *Nanoscale Horizons* **2019**, *4* (2), 434–444. https://doi.org/10.1039/c8nh00274f.

(132) Kim, M.; Lee, J. H.; Nam, J. M. Plasmonic Photothermal Nanoparticles for Biomedical Applications. *Advanced Science* **2019**, *6* (17). https://doi.org/10.1002/advs.201900471.

(133) Zhou, X.; Xu, W.; Liu, G.; Panda, D.; Chen, P. Size-Dependent Catalytic Activity and Dynamics of Gold Nanoparticles at the Single-Molecule Level. *Journal of the American Chemical Society* **2010**, *132* (1), 138–146. https://doi.org/10.1021/ja904307n.

(134) Paquin, F.; Rivnay, J.; Salleo, A.; Stingelin, N.; Silva, C. Multi-Phase Semicrystalline Microstructures Drive Exciton Dissociation in Neat Plastic Semiconductors. *J. Mater. Chem. C* **2015**, *3*, 10715–10722. https://doi.org/10.1039/b000000x.

(135) Chen, T.; Zhang, Y.; Xu, W. Size-Dependent Catalytic Kinetics and Dynamics of Pd Nanocubes: A Single-Particle Study. *Physical Chemistry Chemical Physics* **2016**, *18* (32), 22494–22502. https://doi.org/10.1039/c6cp02719a.

(136) Suchomel, P.; Kvitek, L.; Prucek, R.; Panacek, A.; Halder, A.; Vajda, S.; Zboril, R. Simple Size-Controlled Synthesis of Au Nanoparticles and Their Size-Dependent Catalytic Activity. *Scientific Reports* **2018**, *8* (1), 1–11. https://doi.org/10.1038/s41598-018-22976-5.

(137) Roldan Cuenya, B.; Behafarid, F. Nanocatalysis: Size- and Shape-Dependent Chemisorption and Catalytic Reactivity. *Surface Science Reports* **2015**, *70* (2), 135–187. https://doi.org/10.1016/j.surfrep.2015.01.001.

(138) Cao, S.; Tao, F. F.; Tang, Y.; Li, Y.; Yu, J. Size- and Shape-Dependent Catalytic Performances of Oxidation and Reduction Reactions on Nanocatalysts. *Chemical Society Reviews* **2016**, *45* (17), 4747–4765. https://doi.org/10.1039/c6cs00094k.

(139) Shipway, A. N.; Katz, E.; Willner, I. Nanoparticle Arrays on Surfaces for Electronic, Optical, and Sensor Applications. *Angewandte Chemie (International Edition in English)* **2000**, *39* (15 SUPPL.), 19–52. https://doi.org/10.1002/1439-7641(20000804)1:1<18::aid-cphc18>3.3.co;2-c.

(140) Mafuné, F.; Kohno, J. Y.; Takeda, Y.; Kondow, T.; Sawabe, H. Formation of Gold Nanoparticles by Laser Ablation in Aqueous Solution of Surfactant. *Journal of Physical Chemistry B* **2001**, *105* (22), 5114–5120. https://doi.org/10.1021/jp0037091.

(141) M.C. Daniel, D. A. Gold Nanoparticles: Assembly, Supramolecularchemistry, Quantum-Size-Related Properties, and Applications Toward. *Chem. Rev.* **2004**, *104*, 293–346.

(142) Kimling, J.; Maier, M.; Okenve, B.; Kotaidis, V.; Ballot, H.; Plech, A. Turkevich Method for Gold Nanoparticle Synthesis Revisited. *Journal of Physical Chemistry B* **2006**, *110* (32), 15700–15707. https://doi.org/10.1021/jp061667w.

(143) Pinto Reis, C.; Neufeld, R. J.; Ribeiro, A. J.; Veiga, F. Nanoencapsulation I. Methods for Preparation of Drug-Loaded Polymeric Nanoparticles. *Nanomedicine: Nanotechnology, Biology, and Medicine* **2006**, *2* (1), 8–21. https://doi.org/10.1016/j.nano.2005.12.003.

(144) Kundu, S.; Maheshwari, V.; Niu, S.; Saraf, R. F. Polyelectrolyte Mediated Scalable Synthesis of Highly Stable Silver Nanocubes in Less than a Minute Using Microwave Irradiation. *Nanotechnology* **2008**, *19* (6). https://doi.org/10.1088/0957-4484/19/6/065604.

(145) Satyanarayana, T. □A Review on Chemical and Physical Synthesis Methods of Nanomaterials. *International Journal for Research in Applied Science and Engineering Technology* **2018**, *6* (1), 2885–2889. https://doi.org/10.22214/ijraset.2018.1396.

(146) Turkevich, J.; Stevenson, P. C.; Hillier, J. A Study of the Nucleation and Growth Processes in the Synthesis of Colloidal Gold. *Discussions of the Faraday Society* **1951**, *11* (c), 55–75. https://doi.org/10.1039/DF9511100055.

(147) Yoshimura, H. Protein-Assisted Nanoparticle Synthesis. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **2006**, *282–283*, 464–470. https://doi.org/10.1016/j.colsurfa.2006.01.037.

(148) McDonagh, B. H.; Singh, G.; Bandyopadhyay, S.; Lystvet, S. M.; Ryan, J. A.; Volden, S.; Kim, E.; Sandvig, I.; Sandvig, A.; Glomm, W. R. Controlling the Self-Assembly and Optical Properties of Gold Nanoclusters and Gold Nanoparticles Biomineralized with Bovine Serum Albumin. *RSC Advances* **2015**, *5* (122), 101101–101109. https://doi.org/10.1039/c5ra23423a.

(149) Leng, Y.; Fu, L.; Ye, L.; Li, B.; Xu, X.; Xing, X.; He, J.; Song, Y.; Leng, C.; Guo, Y.; Ji, X.; Lu, Z. Protein-Directed Synthesis of Highly Monodispersed, Spherical Gold Nanoparticles and Their Applications in Multidimensional Sensing. *Scientific Reports* **2016**, *6* (March), 1–11. https://doi.org/10.1038/srep28900.

(150) Hart, C.; Abuladel, N.; Bee, M.; Kreider, M. C.; Cvitan, A. C.; Esson, M. M.; Farag, A.; Ibeh, T.; Kalivas, E. N.; Larco, D. M.; Walker Long, A.; Lymperopoulos, L.; Mendel, Z.; Miles, N.; Zareba, C. M.; Schwabacher, J. C.; Slucher, H.; Vinals, J.; Heddleston, J. M.; Li, W.; Fox, D. M.; Hartings, M. R. Protein-Templated Gold Nanoparticle Synthesis: Protein Organization, Controlled Gold Sequestration, and Unexpected Reaction Products. *Dalton Transactions* **2017**, *46* (47), 16465–16473. https://doi.org/10.1039/c7dt03275g.

(151) Diaz, D.; Care, A.; Sunna, A. Bioengineering Strategies for Protein-Based Nanoparticles. *Genes* **2018**, *9* (7). https://doi.org/10.3390/genes9070370.

(152) Krajina, B. A.; Proctor, A. C.; Schoen, A. P.; Spakowitz, A. J.; Heilshorn, S. C. Biotemplated Synthesis of Inorganic Materials: An Emerging Paradigm for Nanomaterial

Synthesis Inspired by Nature. *Progress in Materials Science* **2018**, *91*, 1–23. https://doi.org/10.1016/j.pmatsci.2017.08.001.

(153) Künzle, M.; Lach, M.; Beck, T. Crystalline Protein Scaffolds as a Defined Environment for the Synthesis of Bioinorganic Materials. *Dalton Transactions* **2018**, *47* (31), 10382–10387. https://doi.org/10.1039/c8dt01192c.

(154) Garcia-Hernandez, C.; Freese, A. K.; Rodriguez-Mendez, M. L.; Wanekaya, A. K. In Situ Synthesis, Stabilization and Activity of Protein-Modified Gold Nanoparticles for Biological Applications. *Biomaterials Science* **2019**, *7* (6), 2511–2519. https://doi.org/10.1039/c9bm00129h.

(155) Yang, M.; Song, W. J. Diverse Protein Assembly Driven by Metal and Chelating Amino Acids with Selectivity and Tunability. *Nature Communications* **2019**, *10* (1), 1–11. https://doi.org/10.1038/s41467-019-13491-w.

(156) Dujardin, E.; Peet, C.; Stubbs, G.; Culver, J. N.; Mann, S. Organization of Metallic Nanoparticles Using Tobacco Mosaic Virus Templates. *Nano Letters* **2003**, *3* (3), 413–417. https://doi.org/10.1021/nl034004o.

(157) Zare, D.; Akbarzadeh, A.; Bararpour, N. Synthesis and Functionalization of Gold Nanoparticles by Using of Poly Functional Amino Acids. *International Journal of Nanoscience and Nanotechnology* **2010**, *6* (4), 223–230.

(158) Zare, D.; Khoshnevisan, K.; Barkhi, M.; Tahami, H. V. Fabrication of Capped Gold Nanoparticles by Using Various Amino Acids. *Journal of Experimental Nanoscience* **2014**, *9* (9), 957–965. https://doi.org/10.1080/17458080.2012.752582.

(159) Kogan, M. J.; Olmedo, I.; Hosta, L.; Guerrero, A. R.; Cruz, L. J.; Albericio, F. Peptides and Metallic Nanoparticles for Biomedical Applications. *Nanomedicine* **2007**, *2* (3), 287–306. https://doi.org/10.2217/17435889.2.3.287.

(160) Yang, T.; Li, Z.; Wang, L.; Guo, C.; Sun, Y. Synthesis, Characterization, and Self-Assembly of Protein Lysozyme Monolayer-Stabilized Gold Nanoparticles. *Langmuir* **2007**, *23* (21), 10533–10538. https://doi.org/10.1021/la701649z.

(161) Li, Y.; Tang, Z.; Prasad, P. N.; Knecht, M. R.; Swihart, M. T. Peptide-Mediated Synthesis of Gold Nanoparticles: Effects of Peptide Sequence and Nature of Binding on Physicochemical Properties. *Nanoscale* **2014**, *6* (6), 3165–3172. https://doi.org/10.1039/c3nr06201e.

(162) Jia, Y.; Yan, X.; Guo, X.; Zhou, G.; Liu, P.; Li, Z. One Step Preparation of Peptide-Coated Gold Nanoparticles with Tunable Size. *Materials* **2019**, *12* (13), 1–12. https://doi.org/10.3390/ma12132107.

(163) JyothiKumar, A.; Kailasnath, M.; Simon, J.; Mathew, S. BSA Stabilized Gold Nanoparticles: Synthesis and Characterization. *Materials Today: Proceedings* **2019**, *9*, 111–115. https://doi.org/10.1016/j.matpr.2019.02.044.

(164) 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*, 0–7. https://doi.org/10.1186/s11671-019-2884-3.

(165) Bari, N. K.; Kumar, G.; Bhatt, A.; Hazra, J. P.; Garg, A.; Ali, M. E.; Sinha, S. Nanoparticle Fabrication on Bacterial Microcompartment Surface for the Development of Hybrid Enzyme-Inorganic Catalyst. *ACS Catalysis* **2018**, *8* (9), 7742–7748. https://doi.org/10.1021/acscatal.8b02322.

(166) Crowley, C. S.; Cascio, D.; Sawaya, M. R.; Kopstein, J. S.; Bobik, T. A.; Yeates, T. O. Structural Insight into the Mechanisms of Transport across the Salmonella Enterica Pdu Microcompartment Shell. *Journal of Biological Chemistry* **2010**, *285* (48), 37838–37846. https://doi.org/10.1074/jbc.M110.160580.

(167) Underwood, S.; Mulvaney, P. Effect of the Solution Refractive Index on the Color of Gold Colloids. *Langmuir* **1994**, *10* (10), 3427–3430. https://doi.org/10.1021/la00022a011.

(168) Link, S.; El-Sayed, M. A. O Ptical P Roperties and U Ltrafast D Ynamics of M Etallic N Anocrystals . *Annual Review of Physical Chemistry* **2003**, *54* (1), 331–366. https://doi.org/10.1146/annurev.physchem.54.011002.103759.

(169) Dintinger, J.; Mühlig, S.; Rockstuhl, C.; Scharf, T. A Bottom-up Approach to Fabricate Optical Metamaterials by Self-Assembled Metallic Nanoparticles. *Optical Materials Express* **2012**, *2* (3), 269. https://doi.org/10.1364/ome.2.000269.

(170) Zuber, A.; Purdey, M.; Schartner, E.; Forbes, C.; Van Der Hoek, B.; Giles, D.; Abell, A.; Monro, T.; Ebendorff-Heidepriem, H. Detection of Gold Nanoparticles with Different Sizes Using Absorption and Fluorescence Based Method. *Sensors and Actuators, B: Chemical* **2016**, *227*, 117–127. https://doi.org/10.1016/j.snb.2015.12.044.

(171) Schrödinger, LLC (2015). The {PyMOL} Molecular Graphics System, Version 2.1.

(172) Pong, B. K.; Elim, H. I.; Chong, J. X.; Ji, W.; Trout, B. L.; Lee, J. Y. New Insights on the Nanoparticle Growth Mechanism in the Citrate Reduction of Gold(III) Salt: Formation of the Au Nanowire Intermediate and Its Nonlinear Optical Properties. *Journal of Physical Chemistry C* **2007**, *111* (17), 6281–6287. https://doi.org/10.1021/jp068666o.

(173) Yau, S. H.; Varnavski, O.; Goodson, T. An Ultrafast Look at Au Nanoclusters. *Accounts of Chemical Research* **2013**, *46* (7), 1506–1516. https://doi.org/10.1021/ar300280w.

(174) Bari, N. K.; Kumar, G.; Hazra, J. P.; 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), 523–533. https://doi.org/10.1039/c9tb02224d.

(175) Lin, C.; Tao, K.; Hua, D.; Ma, Z.; Zhou, S. Size Effect of Gold Nanoparticles in Catalytic Reduction of P-Nitrophenol with NaBH4. *Molecules* **2013**, *18* (10), 12609–12620. https://doi.org/10.3390/molecules181012609.

(176) Aditya, T.; Pal, A.; Pal, T. Nitroarene Reduction: A Trusted Model Reaction to Test Nanoparticle Catalysts. *Chemical Communications* **2015**, *51* (46), 9410–9431. https://doi.org/10.1039/c5cc01131k.

(177) Zhao, P.; Feng, X.; Huang, D.; Yang, G.; Astruc, D. Basic Concepts and Recent Advances in Nitrophenol Reduction by Gold- and Other Transition Metal Nanoparticles. *Coordination Chemistry Reviews* **2015**, *287*, 114–136. https://doi.org/10.1016/j.ccr.2015.01.002.

(178) Gao, L.; Zhuang, J.; Nie, L.; Zhang, J.; Zhang, Y.; Gu, N.; Wang, T.; Feng, J.; Yang, D.; Perrett, S.; Yan, X. Intrinsic Peroxidase-like Activity of Ferromagnetic Nanoparticles. *Nature Nanotechnology* **2007**, *2* (9), 577–583. https://doi.org/10.1038/nnano.2007.260.

(179) Jv, Y.; Li, B.; Cao, R. Positively-Charged Gold Nanoparticles as Peroxidiase Mimic and Their Application in Hydrogen Peroxide and Glucose Detection. *Chemical Communications* **2010**, *46* (42), 8017–8019. https://doi.org/10.1039/c0cc02698k.

(180) Li, R. S.; Liu, H.; Chen, B. Bin; Zhang, H. Z.; Huang, C. Z.; Wang, J. Stable Gold Nanoparticles as a Novel Peroxidase Mimic for Colorimetric Detection of Cysteine. *Analytical Methods* **2016**, *8* (11), 2494–2501. https://doi.org/10.1039/c6ay00367b.

(181) Liu, C. P.; Chen, K. C.; Su, C. F.; Yu, P. Y.; Lee, P. W. Revealing the Active Site of Gold Nanoparticles for the Peroxidase-like Activity: The Determination of Surface Accessibility. *Catalysts* **2019**, *9* (6). https://doi.org/10.3390/catal9060517.

(182) David Josephygg, P.; Elingg, T.; Mason1, R. P. *The Horseradish Peroxidase-Catalyzed Oxidation of 3,5,3',5'-Tetramethylbenzidine FREE RADICAL AND CHARGE-TRANSFER COMPLEX INTERMEDIATES**; 1982; Vol. 257.

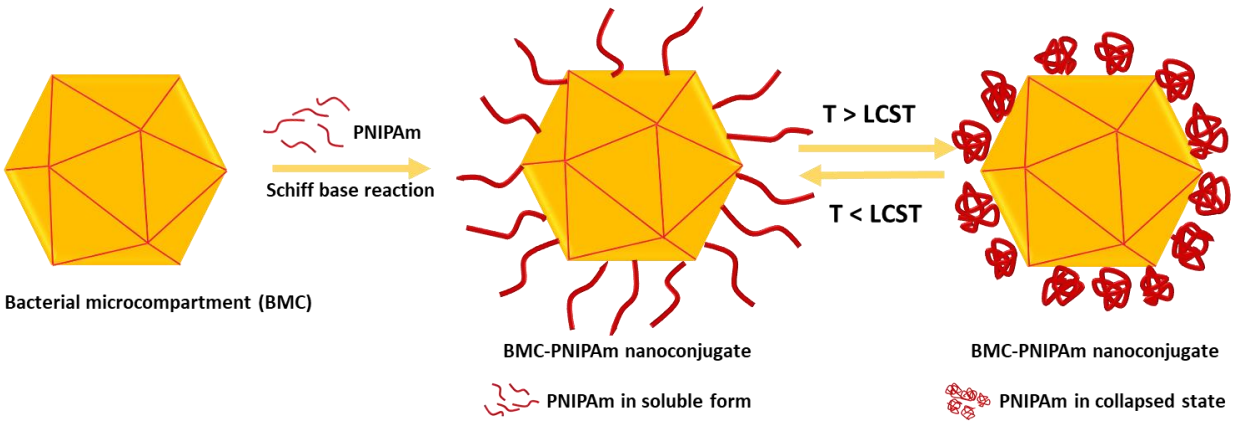
Appendix

Temperature responsive MCP-polymer nanoconjugate

In collaboration with Dr. Raj Kumar Roy, IISER Mohali

Introduction

Bacterial microcompartments are all-proteinaceous self-assembling organelles which are in 100-150nm in size. Because of their proteinaceous shell, they can be targeted via extra functionalization using proteins, antibodies, polymers, etc. They can be used as a synthetic metabolic hub for heterologous cargo loading for various applications like toxic intermediate containment, fast processing of unstable molecules, enhancing catalytic efficiency of slow enzymes by close vicinity, therapeutics, protein stability by preventing proteolytic degradation, and enzymatic activity tuning. Further smart materials can be developed that are actively responsive towards changes in pH, temperature, light, and magnetic field[1–4]. We aimed to design a BMC-polymer nanohybrid by extra functionalization of BMC using a thermoresponsive polymer, PNIPAm (poly-N-isopropyl acrylamide)[5]. PNIPAm is a commonly used thermoresponsive polymer because of its non-toxic nature and is environment friendly[6]. Thermoresponsive polymers exhibit a property of Lower critical solution temperature (LCST) from 30 to 45 ^{O}C[7–9]. LCST is the temperature where the phase transition from hydrophilic to hydrophobic structure occurs. PNIPAm easily dissolves in cold water resulting in a transparent solution, but collapses and aggregates to a globule structure beyond the LCST by excluding water molecules, resulting in a turbid solution. We intended to use this property of thermoresponsive polymer, PNIPAm to tune the catalytic activity of BMC (Scheme 1). Thermoresponsive polymer on attaining LCST temperature will block the pores present in the center of the hexamer and pseudohexamer shell proteins of BMC which are the conduits for substrate channeling and cofactor diffusion. Polymers have been earlier explored to induce the thermal sensitivity to nanoparticles[10–12], enzymes[13–15], proteins[16–20]. BMC architecture is made up of proteins and this provides the advantage of easily conjugating it with polymer. We strategically used aldehyde end-functionalized PNIPAm polymer to conjugate to the amine groups on lysine amino acid present on the shell proteins of BMC using Schiff base reaction. Here, the reaction of an aldehyde with primary amines will form a dynamic covalent imine bond formation (C=N) by elimination of water [Scheme Figure 2(a)] which can be further stabilized by reduction to secondary amines. This approach has been used earlier[21–23].



Scheme 1 Scheme illustrates the hypothesis of bacterial microcompartment being conjugated to the thermoresponsive and aldehyde functionalized polymer PNIPAm. The nanoconjugate formed after conjugation of PNIPAm is shown before and after LCST of PNIPAm in a soluble and collapsed form respectively.

Experimental section

Materials

All materials are procured from Sigma (India). Aldehyde end-functionalized PNIPAm polymers were obtained from Dr. Raj Kumar Roy lab, IISER Mohali.

Methods

LCST determination of PNIPAm

The required concentration of PNIPAm was prepared by dissolving the powder PNIPAm in double distilled water. It was then filtered through Amicon Ultra 0.5ml centrifugal filters with 3kDa MWCO. The characterization of LCST of PNIPAm was performed by measuring the absorbance in circular dichroism (Jasco J-1500, CD spectrophotometer, Jasco, Japan) and size distribution by dynamic light scattering (ZetaSizer Nano ZSP, Malvern Instruments, UK). LCSTs at different concentrations (0.25 mg/ml, 1 mg/ml, 4 mg/ml) of PNIPam were determined by measuring absorbance at 540nm using circular dichroism spectroscopy at 1 °C/min heating from 30 ^{O}C to 42 ^{O}C. The LCST was also determined by size distribution analysis from 30 ^{O}C to 41 ^{O}C by using

dynamic light scattering. The same variation in concentrations was used as used for measuring absorbance.

Differential scanning fluorimetry

To determine the LCST by using Sypro orange (SO) as a probe, the PNIPAm and PNIPam conjugated BMC were incubated with 20x SO concentrations. The incubated mixture was loaded onto 0.2ml thin-96 well PCR plates (BIO-RAD, USA) and sealed with optical quality sealing tape (BIO-RAD, USA). LCST was determined using a CFX96 real-time-PCR instrument (BIO-RAD, USA), and the temperature was ramped from 4 °C to 50 °C at 1°C/cycle with an equilibration time of 60 secs at each temperature. The increase in fluorescence intensity (λex = 450−490 nm, λem = 610−650 nm) was measured by using a CFX96 Touch System, which includes one channel with an LED-filter photodiode combination designated for single-color fluorescence resonance energy transfer (FRET). To study the reversibility of system reaction mixtures were cooled down to 4 °C from 50 °C gradually after the first scan. It was then rescanned to 70 °C under the same experimental conditions. The 1st derivative of the RFU vs temperature plot gave the LCST temperatures.

Purification of BMC (1,2-propanediol utilization microcompartment)

To culture Salmonella enterica LT2, 1% of an overnight grown culture was inoculated in 400 ml of 1X NCE (non-carbon E) medium. The medium was supplemented with 0.6% 1,2-Propanediol, 0.5% succinic acid, and 1 mM of MgSO4. The culture was incubated at 37°C for 16 h and then cells were harvested and washed with buffer A at 8000 X g for 5 minutes (4°C). Buffer A contained 50 mM Tris Base pH 8, 500 mM KCl, 25 mM NaCl, 12.5 mM MgCl2, and 1.5% 1,2-Propanediol. Following this, the cells were 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, and shaken for 30 minutes at room temperature. Then incubated on ice for 5 minutes. The lysed cells were removed by centrifugation at 12,000 x g for 5 minutes at 4°C. BMCs in the supernatant were further pelleted by centrifugation at 20000 x g for 20 min at 4°C. The pellet was

re-suspended in Buffer A (60% of B-PER II and 0.4 mM PMSF) and then centrifuged at 20,000 X g for 20 min at 4 °C. The thin film obtained after discarding the supernatant 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). It was centrifuged at 12,000 X g for 5 min at 4 °C . The supernatant containing PduMCP was stored at 4°C and concentration was estimated using Bradford reagent.

Conjugation of PNIPAm with BMC

For conjugation of aldehyde end-functionalized PNIPAm, the PNIPAm solution was prepared at high concentration and subsequently, the desired concentration was added to a final 500 µl reaction volume (HEPES buffer, pH 5.7). After adding PNIPAm, the BMC at a concentration of 0.25mg/ml was added and incubated at 4 °C for 30 min. Thereafter, sodium cyanoborohydride at 1mM was added and incubated overnight. For removing unconjugated BMC, the mixture was heated to 40°C and centrifuged at 12,000g on an Eppendorf centrifuge for 10 min and the pellet was re-suspended in a buffer. This was repeated twice. Free PNIPam was removed by using centrifugal filters of MWCO 3 kDa.

Characterization and stability of conjugated PNIPAm-BMC

The size analysis of conjugated and non-conjugated BMC was performed by using dynamic light scattering (ZetaSizer Nano ZSP, Malvern Instruments, UK). Conjugation of PNIPAm to BMC was assayed by 12% SDS PAGE (denaturing PAGE) by determining the shifts in protein bands due to conjugation. The stability of conjugated PNIPAm -BMC was analyzed by Fluorescence spectroscopy (FS5 Spectrofluorimeter, Edinburgh Instruments, UK). Temperature-dependent intrinsic fluorescence assay for the enzymatic core (containing aromatic amino acids tyrosine, tryptophan, and phenylalanine) of BMC was performed from 25 °C to 90 °C. 200 µl of conjugated reaction mixture was excited at 280 nm and emission spectra were recorded from 300 to 500 nm. (bandwidth 2nm). Fluorescence intensity at 325 nm was plotted at each temperature for all test conditions.

PNIPAm-BMC Enzymatic Activity assay

The temperature-responsive enzymatic activity of control BMC and PNIPAm conjugated BMC was measured by exploiting the diol dehydratase activity (PduCDE enzyme) present in BMC. Diol dehydratase activity was performed by standard 3-methyl-2-benzothiazoline hydrazine (MBTH) method at different temperatures above and below LCST of PNIPAm. For enzyme assay 5µl BMC and PNIPAm conjugated BMC (0.25mg/ml BMC) were 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). It was incubated for 3 min at a particular temperature condition. 50 µl of 15 µM AdoCbl (adenosylcobalamin) was then added to initiate the reaction and after 10min, 1 ml of citrate buffer (pH 3.6) was added to quench the reaction. 500 µl MBTH of 0.1 % (w/v) was added for 15 min, followed by 1 ml water, and incubated for 10 min. Absorbance was measured at 305 nm using a UV spectrophotometer. To calculate specific activity (product formed in µmol by 1 mg of PduCDE/ min), absorbance was converted to concentration using a molar extinction coefficient 13000 cm^{-1}.

Results and discussion

Characterization of PNIPAm (Poly N-isopropyl acrylamide)

LCST determination of PNIPAm polymer

We initiated the characterization of aldehyde-functionalized PNIPAm (from Dr. Rajkumar Roy, IISER Mohali). We obtained two PNIPAm polymers of different molecular weights, 9kDa and 22kDa, characterized by gel permeation chromatography [Figure 1(a)]. Our next step was to characterize the lower critical solution temperature (LCST) of these polymers. We characterized the LCST temperature using circular dichroism and dynamic light scattering. As thermoresponsive polymers undergo phase transition on heating the polymer chains become insoluble and turn the solution turbid. This phenomenon is exploited while measuring the absorbance of 9kDa and 22kDa PNIPAm from 30 ^{O}C to 42 ^{O}C and absorbance as a function of temperature is plotted [Figure 1(b)]. From the data plots, LCST was found to be dependent on the concentration of the PNIPAm. Different concentrations at 0.25mg/ml, 1mg/ml, 4mg/ml were used and LCST was found to be 40 ^{O}C, 39 ^{O}C, 37 ^{O}C respectively. It is seen that as the concentration of PNIPAm is increased, the LCST is found to be decreasing probably due to a crowded environment[24,25]. It is an essential

property as the conjugation and temperature-dependent tuning of activity of BMC-polymer nanoconjugate requires knowledge about the (LCST) of the polymer.

The LCST was also measured using dynamic light scattering which also provides information about the particle size distribution over a whole temperature range. The size of PNIPAm at 25 ^{O}C was found to be 3.4nm and 5.4 nm as determined from the number mean [Figure 2(c)]. As the PNIPAm collapses from straight conformation with an increase in temperature, its size starts increasing and later it decreases due to the aggregation [Figure 1(c)]. We also analyzed the LCST by using a relatively sensitive and new method, not explored much earlier, where we used a Sypro orange as a probe which is known to bind hydrophobic patches in a protein, and on thermal ramp, the increase in fluorescence allows the determination of their melting temperature. We explored this property to determine the LCST where on incubation of PNIPAm to Sypro orange dye, it increases fluorescence when heated from 4 ^{O}C to 50 ^{O}C [Figure 1(d)]. As expected we found the LCST temperatures obtained from all the methods were comparable. [Table 1].

Preparation of BMC-PNIPAm polymer conjugate

Once we have characterized the LSCT of PNIPAm we moved on to conjugating BMC to PNIPAm polymer. To develop a nanoconjugate using a BMC and a PNIPAm polymer, we conjugated the aldehyde end-functionalized PNIPAm using Schiff base reaction [Figure 2(a)] to the basic amino acids present on the shell protein containing primary amine at pH 5.7[26]. BMC was found to be stable at this low pH condition as determined by the secondary structure which gave a characteristic α-helix spectrum with no significant change [Figure 2(b)]. BMC was used at a concentration of 0.25mg/ml and PNIPAm polymer was used at different concentrations at 0.25mg/ml, 1mg/ml, and 4mg/ml concentration. The reaction mixture of BMC and aldehyde functionalized PNIPAm was prepared in HEPES buffer and incubated at 4 °C on a rocker. HEPES buffer was found to be the most suitable buffer where BMC and PNIPAm were stable and prevented the aggregation of PNIPAm. After 30 min of ongoing reaction, sodium cyanoborohydride was added to reduce unstable imine to a stable amine and kept overnight. The unconjugated BMC was then separated by centrifugation at 40°C and at 12,000g for 10 min and

the unconjugated polymer was separated by using centrifugal filters (MWCO 3 kDa) before proceeding for enzymatic activity

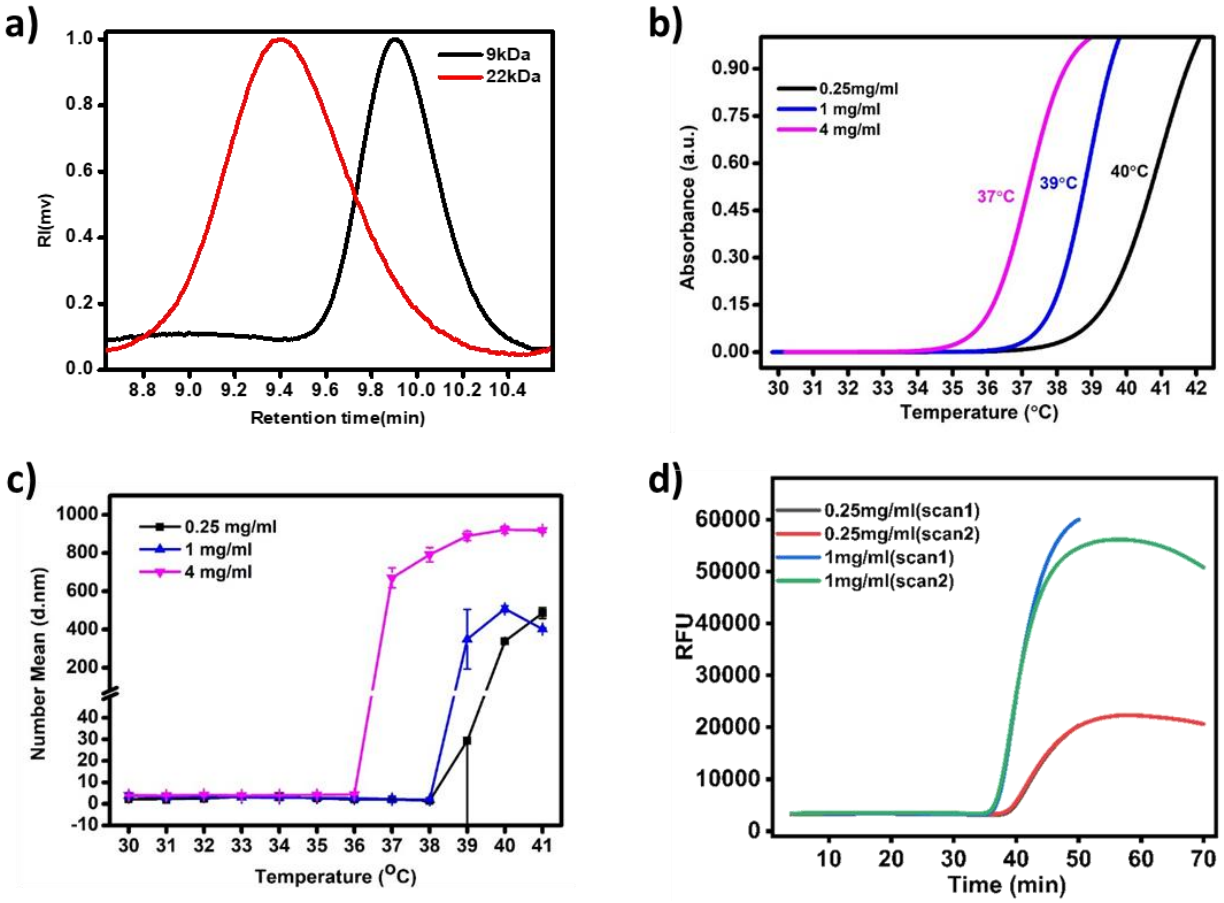


Figure 1 a) Molecular weight determination of PNIPAm polymer using gel permeation chromatography b) Determination of Lower Critical Solution Temperature (LCST) of PNIPAm of different sizes by measuring absorbance from 30 ^{O}C to 42 ^{O}C using circular dichroism c) Temperature-dependent size distribution of PNIPAm from 30 ^{O}C to 41 ^{O}C for LCST determination using dynamic light scattering d) LCST measurement using sypro orange as a probe from 4 ^{O}C to 50 ^{O}C and rescan to 70 ^{O}C.

Characterization of BMC-PNIPAm polymer conjugate

The prepared BMC-PNIPAm nanoconjugates were characterized by studying the differences in the size of the modified BMC. The hydrodynamic diameter changed after conjugation with PNIPAm measured at 25 °C. The hydrodynamic diameter of BMC alone is measured as 119.7 ± 4.4. After conjugation, it is found to increase to 121.8 ± 1.3 and 127.2 ± 2.9 with 9kDa and 22kDa respectively. The increase in size corresponds to the size of PNIPAm alone [Figure 2 (c,d)]. Next, we performed SDS-PAGE with PNIPAm conjugated and non-conjugated BMC. In denaturing

Table 1 LCST determination of PNIPAm by measuring absorbance in circular dichroism, size distribution by dynamic light scattering, and by using Sypro orange probe by differential scanning fluorimetry.

	LCST (Absorbance)	LCST (Dynamic light scattering)	LCST (Differential scanning Fluorimetery)
PNIPAm (9kDa) (0.25mg/ml)	40.5 ^{0}C	40.2 ^{0}C	40.7
PNIPAm (9kDa) (1mg/ml)	38.9 ^{0}C	38.2 ^{0}C	38.5
PNIPAm (9kDa) (4mg/ml)	37.1 ^{0}C	37 ^{0}C	-
PNIPAm (22kDa) (0.25mg/ml)	37.5 ^{0}C	37.2 ^{0}C	36.8
PNIPAm (22kDa) (1mg/ml)	35.9 ^{0}C	35.1 ^{0}C	34.9
PNIPAm (22kDa) (4mg/ml)	34.5 ^{0}C	33.9 ^{0}C	-

PAGE analysis, we expected to observe some high molecular weight band shift. [Figure 2 (e)] shows the presence of high molecular weight bands and the absence of a few low molecular weight bands of 9kDa PNIPAm conjugation at 1mg/ml and 4mg/ml concentration. These results suggest that the PNIPAm has been conjugated to the BMC surface. Further, to understand the stability of the nanoconjugates, we performed an intrinsic fluorescence assay of BMC alone and the conjugated BMC. BMC has an enzymatic core where enzymes possess the aromatic amino acids (tryptophan, tyrosine, phenylalanine) which are responsible for fluorescence when excited. When the BMC shell is intact it gives an emission fluorescence at 325nm. We measured the emission fluorescence at a temperature range from 25 ^{O}C to 45 ^{O}C and plotted for different 9kDa PNIPAm concentrations conjugated to BMC [Figure 3 (a)]. We found a concentration-dependent (PNIPA) decrease in fluorescence intensity of nanoconjugates when compared to BMC alone which suggests the polymer at high concentration destabilizes the BMC, disrupts the shell, and might have diffused into the BMC and denatured the core. There is also a significant decrease in fluorescence at high temperatures above 35 ^{O}C and the effect was enhanced on conjugating the high PNIPAm concentration which points towards the destabilization of the system.

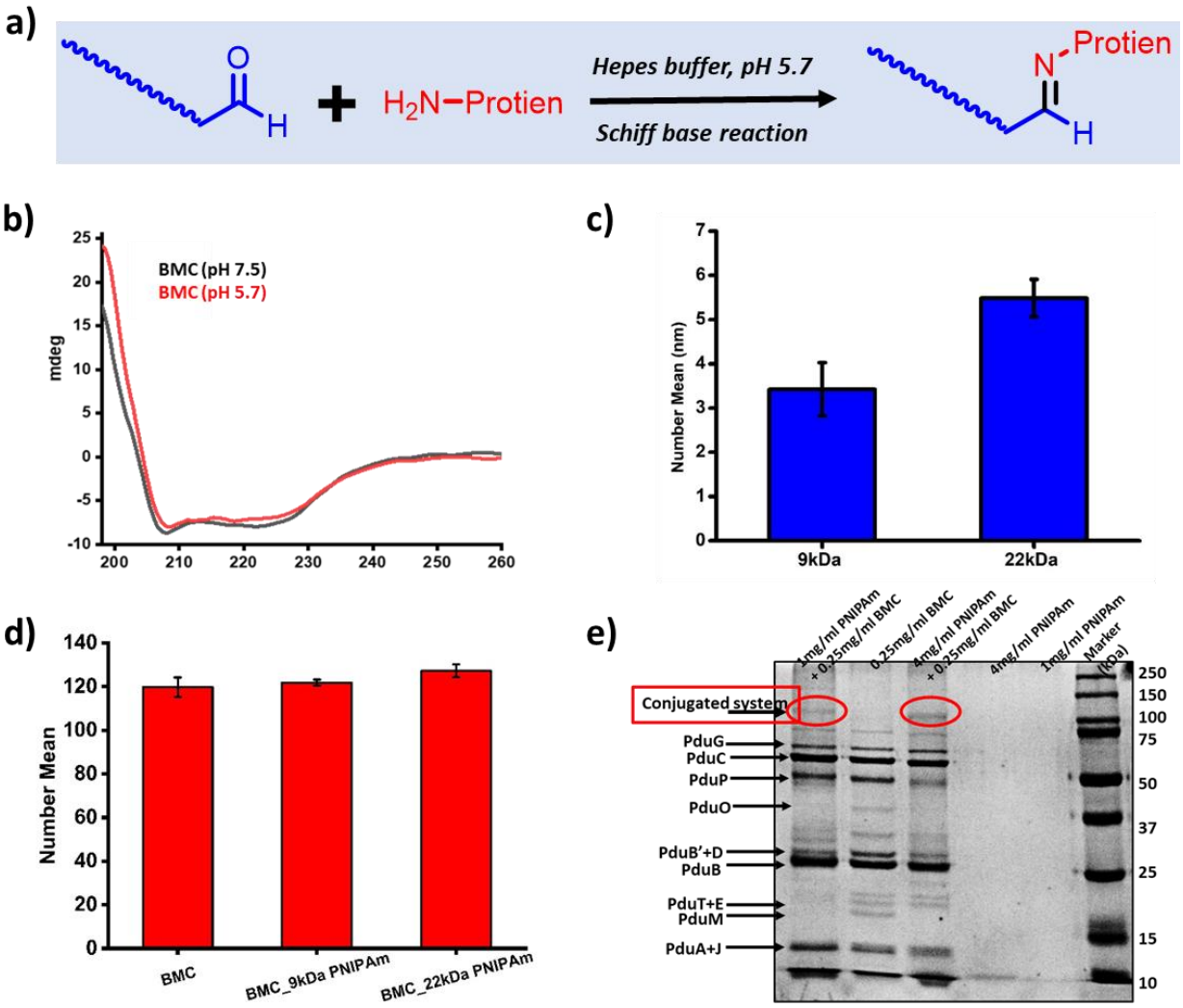


Figure 2 a) Scheme for representing Schiff base reaction for conjugating aldehyde end-functionalized PNIPAm to BMC shell surface b) circular dichroism spectra showing stability of BMC at low pH 5.7 c) Dynamic light scattering showing the size distribution of PNIPAm from number mean data d) Dynamic light scattering showing the difference in size of BMC after conjugation to PNIPAm e) SDS PAGE showing the presence of high molecular weight bands after conjugation and absent in BMC alone.

Enzymatic activity of BMC-PNIPAm polymer conjugate

To investigate whether conjugation of a thermoresponsive PNIPAm can regulate the activity of BMC, we proceeded with the diol dehydratase activity of PduCDE, the major enzyme presents in the BMC [Figure 3 (b)]. PduCDE activity is dependent on cofactor vitamin B_{12} (adenosylcobalamin) and is active against 1,2 propanediol substrate which is suggested to diffuse into the microcompartment through pores present in shell proteins. While performing the activity assay the cofactor and substrate are added into the reaction buffer. Here, depending on the PNIPAm state they are allowed or restricted to diffuse into the BMC. Once the catalysis occurs the product propionaldehyde is detected and measured spectroscopically at 305nm. The BMC is catalytically active till 50 C and the LCST of 9kDa and 22kDa PNIPAm at different concentrations were found to be in the range of 34 ^{O}C to 42 ^{O}C. So we performed the activity assay at 16 ^{O}C, 30 ^{O}C, 37 ^{O}C, and 42 ^{O}C [Figure 3 (c,d)]. According to our hypothesis, we expected a decrease in BMC activity below LCST of the PNIPAm because the elongated and hydrophilic state of PNIPAm will sterically create a layer where the substrate and cofactor were not able to diffuse into the BMC through covered shell protein pores. However, above LCST when PNIPAm chains collapse and open access to the pores to increase the activity again or can block the pores to block the activity. We could not, however, correlate the temperature-dependent change in activity around PNIPAm LCST significantly. Our observation was that the BMC activity was decreased at high molecular weight PNIPAm concentrations significantly after 1mg/ml which can be correlated to the stability data from fluorescence where decreased fluorescence intensity indicates the destabilization of the system.

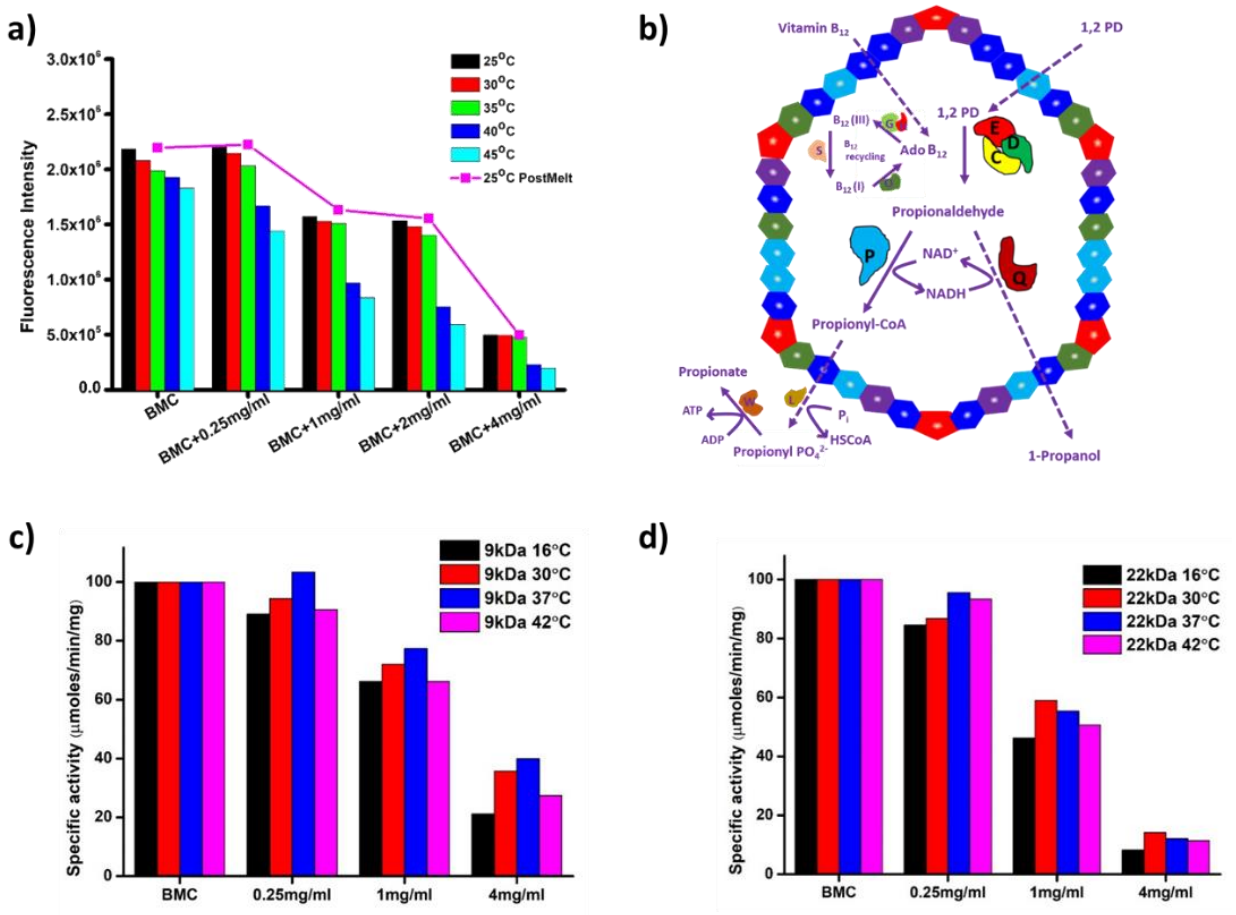


Figure 3 a) Temperature-dependent fluorescence of PNIPAm-BMC conjugate to study the effect of conjugation on BMC stability b) Schematic representation of Pdu BMC c) Temperature dependant activity of PNIPAm-BMC conjugate with 9kDa PNIPAm and d) 22kDa PNIPAm

Conclusions

In this study, we have successfully conjugated the aldehyde functionalized PNIPAm polymer to BMC shell to develop a thermoresponsive nanocontainer using a Schiff base reaction. Interestingly, we observed a concentration-dependent decrease in activity and a small but significant effect was observed on temperature-dependent activity. To increase the yield of the conjugated product and obtain an active temperature-responsive activity of the nanoconjugate, we

have explored two different molecular weight of the polymer (PNIPAm) at different concentrations. Although it is difficult to determine the exact number and location of polymer attachment sites onto the BMC shell but the decrease in activity is quite significant for higher molecular weight polymer. In summary, we used thermoresponsive PNIPAm polymer to develop a tunable nanoreactor out of bacterial microcompartment. We consider that results generated here might help in further standardization, and analysis, and open up new avenues for producing a minimalistic nanoreactor with great potential in various applications.

References

(1) Browne, W. R.; Feringa, B. L. Making Molecular Machines Work. *Nature Nanotech* **2006**, *1* (1), 25–35. https://doi.org/10.1038/nnano.2006.45.

(2) Motornov, M.; Roiter, Y.; Tokarev, I.; Minko, S. Stimuli-Responsive Nanoparticles, Nanogels and Capsules for Integrated Multifunctional Intelligent Systems. *Progress in Polymer Science* **2010**, *1–2* (35), 174–211. https://doi.org/10.1016/j.progpolymsci.2009.10.004.

(3) Zubair Iqbal, M.; Ali, I.; Khan, W. S.; Kong, X.; Dempsey, E. Reversible Self-Assembly of Gold Nanoparticles in Response to External Stimuli. *Materials & Design* **2021**, *205*, 109694. https://doi.org/10.1016/j.matdes.2021.109694.

(4) Shakya, A. K.; Nandakumar, K. S. An Update on Smart Biocatalysts for Industrial and Biomedical Applications. *J R Soc Interface* **2018**, *15* (139), 20180062. https://doi.org/10.1098/rsif.2018.0062.

(5) Schild, H. G. Poly(N-Isopropylacrylamide): Experiment, Theory and Application. *Progress in Polymer Science* **1992**, *17* (2), 163–249. https://doi.org/10.1016/0079-6700(92)90023-R.

(6) Shaibie, N. A.; Ramli, N. A.; Mohammad Faizal, N. D. F.; Srichana, T.; Mohd Amin, M. C. I. Poly(N-Isopropylacrylamide)-Based Polymers: Recent Overview for the Development of Temperature-Responsive Drug Delivery and Biomedical Applications. *Macromolecular Chemistry and Physics n/a* (n/a), 2300157. https://doi.org/10.1002/macp.202300157.

(7) Heskins, M.; Guillet, J. E. Solution Properties of Poly(N-Isopropylacrylamide). *Journal of Macromolecular Science: Part A - Chemistry* **1968**, *2* (8), 1441–1455. https://doi.org/10.1080/10601326808051910.

(8) Zhang, G.; Wu, C. Folding and Formation of Mesoglobules in Dilute Copolymer Solutions. In *Conformation-Dependent Design of Sequences in Copolymers I*; Khokhlov, A. R., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2006; Vol. 195, pp 101–176. https://doi.org/10.1007/12_050.

(9) Halperin, A.; Kröger, M.; Winnik, F. M. Poly(N-Isopropylacrylamide) Phase Diagrams: Fifty Years of Research. *Angewandte Chemie International Edition* **2015**, *54* (51), 15342–15367. https://doi.org/10.1002/anie.201506663.

(10) Durand-Gasselin, C.; Sanson, N.; Lequeux, N. Reversible Controlled Assembly of Thermosensitive Polymer-Coated Gold Nanoparticles. *Langmuir* **2011**, *27* (20), 12329–12335. https://doi.org/10.1021/la2023852.

(11) Li, C.; Wang, C.; Ji, Z.; Jiang, N.; Lin, W.; Li, D. Synthesis of Thiol-Terminated Thermoresponsive Polymers and Their Enhancement Effect on Optical Limiting Property of Gold Nanoparticles. *European Polymer Journal* **2019**, *113*, 404–410. https://doi.org/10.1016/j.eurpolymj.2019.02.009.

(12) Pereira, S. O.; Barros-Timmons, A.; Trindade, T. Polymer@gold Nanoparticles Prepared via RAFT Polymerization for Opto-Biodetection. *Polymers* **2018**, *10* (2), 189. https://doi.org/10.3390/polym10020189.

(13) Hijazi, M.; Türkmen, E.; Tiller, J. C. Full Thermal Switching of Enzymes by Thermoresponsive Poly(2-Oxazoline)-Based Enzyme Inhibitors. *Chemistry – A European Journal* **2020**, *26* (59), 13367–13371. https://doi.org/10.1002/chem.202001909.

(14) Shimoboji, T.; Larenas, E.; Fowler, T.; Hoffman, A. S.; Stayton, P. S. Temperature-Induced Switching of Enzyme Activity with Smart Polymer−Enzyme Conjugates. *Bioconjugate Chem.* **2003**, *14* (3), 517–525. https://doi.org/10.1021/bc025615v.

(15) Heredero, M.; Beloqui, A. Enzyme-Polymer Conjugates for Tuning, Enhancing, and Expanding Biocatalytic Activity. *ChemBioChem* **2023**, *24* (4), e202200611. https://doi.org/10.1002/cbic.202200611.

(16) Stayton, P. S.; Shimoboji, T.; Long, C.; Chilkoti, A.; Ghen, G.; Harris, J. M.; Hoffman, A. S. Control of Protein–Ligand Recognition Using a Stimuli-Responsive Polymer. *Nature* **1995**, *378* (6556), 472–474. https://doi.org/10.1038/378472a0.

(17) Hoffman, A. S. Bioconjugates of Intelligent Polymers and Recognition Proteins for Use in Diagnostics and Affinity Separations. *Clinical Chemistry* **2000**, *46* (9), 1478–1486. https://doi.org/10.1093/clinchem/46.9.1478.

(18) Broyer, R. M.; Grover, G. N.; Maynard, H. D. Emerging Synthetic Approaches for Protein-Polymer Conjugations. *Chem Commun (Camb)* **2011**, *47* (8), 2212–2226. https://doi.org/10.1039/c0cc04062b.

(19) Heredia, K. L.; Maynard, H. D. Synthesis of Protein–Polymer Conjugates. *Org. Biomol. Chem.* **2006**, *5* (1), 45–53. https://doi.org/10.1039/B612355D.

(20) Trzebicka, B.; Szweda, R.; Kosowski, D.; Szweda, D.; Otulakowski, Ł.; Haladjova, E.; Dworak, A. Thermoresponsive Polymer-Peptide/Protein Conjugates. *Progress in Polymer Science* **2017**, *68*, 35–76. https://doi.org/10.1016/j.progpolymsci.2016.12.004.

(21) Sahajpal, K.; Shekhar, S.; Kumar, A.; Sharma, B.; Meena, M. K.; Bhagi, A. K.; Sharma, S. Dynamic Protein and Polypeptide Hydrogels Based on Schiff Base Co-Assembly for Biomedicine. *J. Mater. Chem. B* **2022**, *10* (17), 3173–3198. https://doi.org/10.1039/D2TB00077F.

(22) Abdelaty, M. S. A.; Kuckling, D. Poly (N-Isopropyl Acrylamide-Co-Vanillin Acrylate) Dual Responsive Functional Copolymers for Grafting Biomolecules by Schiff's Base Click Reaction. *Open Journal of Organic Polymer Materials* **2018**, *8* (2), 15–32. https://doi.org/10.4236/ojopm.2018.82002.

(23) Hamley, I. W. PEG–Peptide Conjugates. *Biomacromolecules* **2014**, *15* (5), 1543–1559. https://doi.org/10.1021/bm500246w.

(24) Plunkett, K. N.; Zhu, X.; Moore, J. S.; Leckband, D. E. PNIPAM Chain Collapse Depends on the Molecular Weight and Grafting Density. *Langmuir* **2006**, *22* (9), 4259–4266. https://doi.org/10.1021/la0531502.

(25) Pham, Q.-T.; Yao, Z.-H.; Chang, Y.-T.; Wang, F.-M.; Chern, C.-S. LCST Phase Transition Kinetics of Aqueous Poly(N-Isopropylacrylamide) Solution. *Journal of the Taiwan Institute of Chemical Engineers* **2018**, *93*, 63–69. https://doi.org/10.1016/j.jtice.2018.07.045.

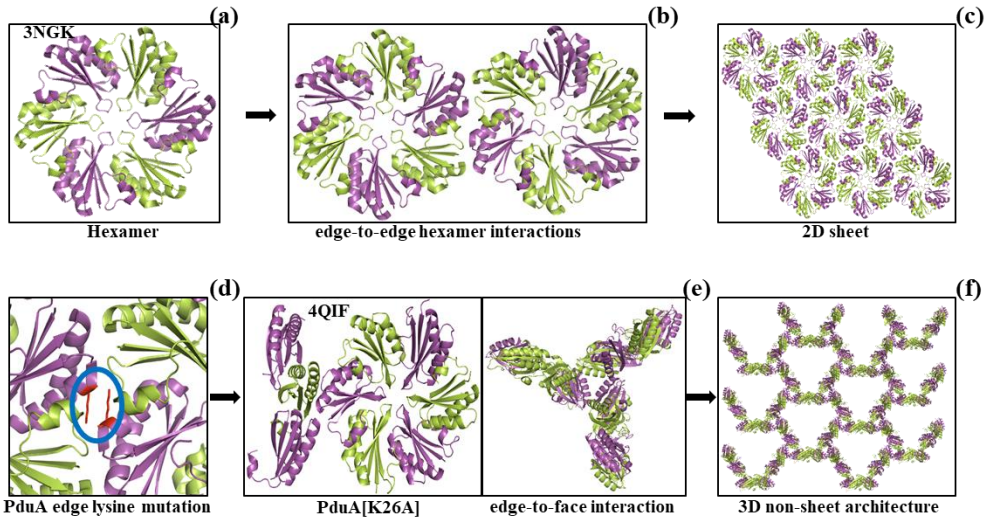
(26) Liu, Y.; Lee, J.; Mansfield, K. M.; Ko, J. H.; Sallam, S.; Wesdemiotis, C.; Maynard, H. D. Trehalose Glycopolymer Enhances Both Solution Stability and Pharmacokinetics of a Therapeutic Protein. *Bioconjugate Chem.* **2017**, *28* (3), 836–845. https://doi.org/10.1021/acs.bioconjchem.6b00659.

List of Publications from thesis

Kaur, S.; Bari, N. K.; Sinha, S. Varying Protein Architectures in 3-Dimensions for Scaffolding and Modulating Properties of Catalytic Gold Nanoparticles. *Amino Acids* **2022**, *54* (3), 441–454. https://doi.org/10.1007/s00726-022-03127-7.

Abstract

Fabrication and development of nanoscale materials with tunable structural and functional properties require a dynamic arrangement of nanoparticles on architectural templates. The function of nanoparticles not only depends on the property of the nanoparticles but also on their spatial orientations. Proteins with self-assembling properties which can be genetically engineered to varying architectural designs for scaffolds can be used to develop different orientations of nanoparticles in three dimensions. Here, we report the use of naturally self-assembling bacterial micro-compartment shell protein (PduA) assemblies in 2D and its single-point mutant variant (PduA[K26A]) in 3D architectures for the reduction and fabrication of gold nanoparticles. Interestingly, the different spatial organization of gold nanoparticles resulted in a smaller size in the 3D architect scaffold. Here, we observed a two-fold increase in catalytic activity and six-fold higher affinity toward TMB (3,3',5,5'-tetramethylbenzidine) substrate as a measure of higher peroxidase activity (nanozymatic) in the case of PduA[K26A] 3D scaffold. This approach demonstrates that the hierarchical organization of scaffold enables the fine-tuning of nanoparticle properties, thus paving the way toward the design of new nanoscale materials.



List of Publications outside thesis

1) Naimat K. Bari, Gaurav Kumar, Jagadish P. Hazra, **Simerpreet Kaur**, Sharmistha Sinha. (2020); J Mater Chem B. Functional protein shells fabricated from the self-assembling protein sheets of prokaryotic organelles. 8 (3): 523-533.
2) Naimat K. Bari, Jagadish P. Hazra, Gaurav Kumar, **Simerpreet Kaur**, Sharmistha Sinha. (2020); Biochimica et Biophysica acta. General Subjects. Probe into a multi-protein prokaryotic organelle using thermal scanning assay reveals distinct properties of the core and the shell. 1864(10): 129680.
3) Jagadish P. Hazra, Amin Sagar, Nisha Arora, Debadutta Deb, **Simerpreet Kaur**, and Sabyasachi Rakshit. (2019); BioChem. J. Broken force dispersal network in tip-links by the mutations at the Ca2+ -binding residues induces hearing-Loss. 476 (16): 2411-2425.