Investigating the Structure, Function and Dynamics of an Intrinsically Disordered Enzyme

A Thesis Submitted in Partial Fulfilment of the
Requirements for the Degree of

Doctor of Philosophy

By
Ms. Shrutidhara Biswas
Roll no. 11610609

Department of Biosciences and Bioengineering
Indian Institute of Technology Guwahati
Guwahati, Assam- 781039, India

February 2020

INDIAN INSTITUTE OF TECHNOLOGY
GUWAHATI, Assam, India

Department of Biosciences and Bioengineering

STATEMENT

I do hereby declare that the matter embodied in this thesis is the result of investigations carried out by me in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India, under the guidance of Prof. Rajaram Swaminathan.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators.

IIT Guwahati

February, 2020

Shrutidhara Biswas

Roll no. 11610609

INDIAN INSTITUTE OF TECHNOLOGY
GUWAHATI, Assam, India

Department of Biosciences and Bioengineering

It is certified that the work described in this thesis, entitled "***Investigating the structure, function and dynamics of an intrinsically disordered enzyme***" done by **Ms. Shrutidhara Biswas** for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India, and this work has not been submitted elsewhere for a degree.

IIT Guwahati

February, 2020

Prof. Rajaram Swaminathan
Dept. of Biosciences and Bioengineering
IIT Guwahati

Dedicated to

my

family and friends

Acknowledgements

Pursuing a Ph.D. degree has been a great learning lesson for me, both from a professional and personal prospective. I believe all the challenges I encountered during this extensive process, be it in conducting experiments or while interacting with colleagues and superiors, has made me a more patient, sensible and humble person. I now understand and have come to highly appreciate the kind of endurance, passion, sacrifice and sheer will-power that scientists throughout the course of human civilization must have possessed that made possible for the present generation of humans to enjoy a life of great convenience and ease.

I take this opportunity to give credit and my heartfelt gratitude to all those who have made it possible for me to begin, give direction and finally bring my Ph.D. work to completion.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. R. Swaminathan for allowing me to work under his guidance and providing me with a conducive environment to carry out my research. He believed in me and gave me a lot of liberty to address scientific problems as I deemed fit. His curious mind and passion for good science will always be a source of inspiration for me. He has been an excellent teacher, from whom I have learnt a lot of things. Without his support and patience, my research work would not have been brought to its completion.

I would like to extend my sincere thanks to my doctoral committee chairperson, Prof S.S. Ghosh, for his constant support and encouragement during my Ph.D. tenure. His critical insights and valuable suggestions have helped me give proper direction to my research work. He has always tried to make my professional life at IIT Guwahati hassle-free and has generously given me access to his lab instruments and resources. Moreover, his counselling and trust in my capabilities has given me the much-needed strength and confidence to finish my work and move forward in life.

I am also grateful to my doctoral committee members, Dr. Bhubaneswar Mandal and Dr. Debasish Das for their valuable suggestions which helped me to widen my research from various perspectives.

I am indebted to Dr. P. Srihari and P. Gangadhar, IICT, Hyderabad for investing their time and energy in the challenging process of the compound (TSA, ligand) synthesis which was a vital step for the project I undertook.

I also thank Prof. D. Hilvert, ETH Zurich, Switzerland for kindly gifting us the plasmid DNA (mMjCM) and providing necessary background information to work on monomeric chorismate mutase (mCM).

My Ph.D. journey was very enjoyable with the love and support of my lab mates and friends Vijya, Yoganand, Himanshu, Saumya, Tulsi, Zia, Amrendra, Dileep, Anurag, Ekramul, Bholanath, Shubham, Aditya, Garima, Shubhojit, Abheek, Alka, Nujelu and Hiramoni. I made a family for life at IIT Guwahati and will cherish my Ph.D. days forever.

My deepest appreciation is for my parents who stood my me through thick and thin during this entire period. The continuous emotional and financial support they provided me let me concentrate on my work at hand without any other worries. Their believe that I can achieve something of substance through my Ph.D. gave me the will to push through many challenges. My younger brother Abhishek has always been my side whenever I needed him. His love and support have been a great bonus for me. I express my sincere gratitude to my in-laws who have been extremely supportive to carry forward my studies. I indeed consider myself very lucky and blessed to have such encouraging family members.

Although it is not possible to name here each and every person who helped me through this journey, as there were many, I would nevertheless like to thank each one of them for helping me in some way or other during different stages of my Ph.D.

I am grateful to my dear husband, Dr. Anand Kant Das, who is my support system and mentor. His understanding, patience and unconditional love encouraged me to carry forward my work and not give up. He has been a source of true inspiration for me as a human being and a scientific researcher. He has moved mountains to make me a happy and confident person. I owe every bit of my achievement to my family.

Last but not the least, I thank the Almighty who I believe has always protected me, heard my prayers in times of dire need and helped me make the right choices in life.

Table of contents

List of figures

Figure 1.8. Measurements of predicted disorder in reference proteomes and measurements of missing regions in the PDB. Box plots are first quartile, second quartile (median), and third quartile. Plot whiskers extend to 5% and 95% of the data. (a and b) The fraction of predicted disorder per-protein in reference proteomes in eukaryotes (a) and bacteria (b). (c and d) The fraction of missing residues per-protein in the PDB in eukaryotes (c) and bacteria (d). (a–d) The blue diamond shows the relative fraction of the full set of enzymes that each enzyme type occupies. (e and f) The longest predicted missing region per-protein in reference proteomes in eukaryotes (e) and bacteria (f). The red star shows the average protein length. (g and h) The longest missing region per-protein in eukaryotes (g) and bacteria (h). The red diamond shows the average number of removed residues per-protein before crystallization.

Figure 1.9. FRET (fluorescence resonance energy transfer) within a protein molecule. D, donor; A, acceptor. Zigzag arrows of violet, light blue, and orange indicate excitation light for the donor, emission from the donor, and emission from the acceptor, respectively.

Figure 1.10. Turkey skeletal 2Ca–Troponin C. *(Left)* The crystal structure *(Right)* Tryptophan-22 fluorescence lifetime intensity decays of Troponin C without acceptor (D) and with an acceptor (DA), an IAEDANS on cysteine 52. *Top panel*: without Ca^{2+} and *Bottom panel*: with Ca^{2+}.

Figure 1.11. Topological redesign of the enzyme MjCM. (*Left*) The thermostable MjCM homodimer (Left) was converted into a monomer (mMjCM, middle) by inserting a flexible hinge loop (red) into the long H1 helix. The models are based on the X-ray structure of a related *E. coli* chorismate mutase domain complexed with a transition-state analog (TSA), which is shown in in ball-and-stick representation in the model. (*Right*) Both enzymes efficiently catalyze the rearrangement of chorismate to prephenate.

Figure 1.12. Active site of monomeric chorismate mutase (sequence numbering based on EcCM) [Adapted from Vamvaca et al.2008].

Figure 2.1. Jablonski diagram [Source: (Lakowicz, 2006)]

Figure 2.2. Image above showing the schematic of fluorescence anisotropy (Source: Principle in Fluorescence Spectroscopy by J. R. Lakowicz, third edition, 2006).

Figure 2.3. Schematic for Time-domain fluorescence lifetime measurements (Adapted from Principle in Fluorescence Spectroscopy by J. R. Lakowicz, third edition, 2006).

Figure 2.4. Illustration of the global and localized (Trp probe) motion in the protein mCM.

Figure 3.1. Graphical representation of the NMR structure of mCM (PDB: 2gtv) bound to the ligand (TSA) showcasing the sites of mutation and FRET distances. The respective distances from N-atom of indole in Trp 26 to S-atom of Cys 69 and C-atom of all the selected alanine residues (Ala 6, Ala 32 and Ala 92) in the ligand-bound NMR structure are shown in dashed red lines and numbers. The sequence of wild-type mCM is also shown and the sites planned for mutations highlighted in red

and numbered. UCSF Chimera 1.13.1 visualization and analysis tool was used for the distance calculations and graphical representation of mCM.

Figure 3.2. Strategy used for creation of mutation at a single site for amino acid replacement through Site-Directed mutagenesis. Parental plasmid DNA represented as green and yellow concentric circles. Primers and newly synthesized daughter DNA plasmid represented as pink and blue concentric circles. The primers are represented with a arrow head and the site of mutation marked as a small cross in the strands/circles. The digested parent DNA is reprerstented by hyped circles.

Figure 3.3. *(Left)* Agarose gel showing mutated plasmid mMjCM DNA from positive colonies (1,2,3,4,5) of Mutant 1 (W24K). Parental digested DNA (D) fragments after *Dpn I* digestion can be observed in the gel. *(Right)* Luria-broth agar plates containing ampicillin showing positive colonies of Mutant 1 after PCR-like amplification and transformation into DH5α bacterial cells.

Figure 3.4. 15% SDS PAGE gel showing expression of mCM wild-type protein. Some protein can also be observed in the pellet (lane 4).

Figure 3.5. 15% SDS PAGE gel showing expression and purification of mCM wild type protein and its mutants (A)wild-type mCM, (B)Mutant 1, (C)Mutant 2, (D) Mutant 3

Figure 3.6. The mass spectrum (MALDI) of Wt mCM. Observed mass was 13174 Da.

Figure 3.7. The mass spectrum (MALDI) of Mutant1. Observed mass was 13104 Da.

Figure 3.8. The mass spectrum (MALDI) of Mutant2. Observed mass was 13071 Da.

Figure 3.9. The mass spectrum (MALDI) of Mutant3. Observed mass was 13102 Da.

Figure 3.10. Simplified 2-D representation of Wt mCM and its three mutants (Mutant1, Mutant2 and Mutant3) that were expressed and purified for further biochemical and biophysical investigations. The mutated sites in the individual mutants are shown as red crosses. Tryphophan sites are marked as green diamonds. A dansyl probe labeled to the single cysteine for fluorescence studies in mCM is shown as light orange hexagon.

Figure 4.1. Normalized absorbance spectra of mCM and its mutants. [Inset: protein absorbance between 270-325 nm with a wavelength maximum ranging from 274-280 nm.]

Figure 4.2. Fluorescence spectra (steady-state) of mCM and its mutants obtained after excitation with 295 nm wavelength (excitation slit: 2 nm, emission slit: 5 nm). 20 µM of each of the proteins were used for the experiment. Fluorescence emission of NATA (10 µM) is shown as dotted curve.

Figure 4.3. Fluorescence spectra of 10 µM Wt mCM (unlabeled with any probe) in presence of 20 µM (1:2) and 200 µM (1:20) TSA.

Figure 4.4. (*Top*) The binding of ANS (40µM) to Wt mCM and all the mutants (20µM). Addition of TSA (40 µM) to mCM (20 µM) before ANS binding affected the binding process negatively. An excitation slit of 2 nm and emission slit of 10 nm was used for the measurements and all samples were excited at 375 nm. (*Bottom, left*) Integrated fluorescence area and (*Bottom, right*) emission λ_{max} wavelengths corresponding to the ANS fluorescence intensities for Wt mCM and its mutants in absence and presence of TSA.

Figure 4.5. Steady state FRET between Trp24/26 and Cys 69 in Wt mCM. Emission spectra corresponding to both Trp and the dansyl probe are observed after excitation at 295 nm (solid lines). Trp in unlabeled mCM is represented in coloured dashed lines. Addition of GdnHCl gradually increases Trp fluorescence in labeled mCM, whereas the fluorescence of unlabeled mCM remains unaffected in presence of Gdn.HCl.

Figure 4.6. NMR structure of Wt mCM focusing on the Trp 24 and Trp 26 residues in a ball and stick representation (red). The distance between the two tryptophan (NE1:NE1) is calculated to be approximately 8.28 Å (VMD program).

Figure 4.7. CD spectra of mCM and its mutants. A concentration of 15 µM protein in PBS pH 7.2 buffer was used for the CD measurement. (Inset) The values of the two characteristics peaks of α-helices in CD spectra are compared across mCM and its mutants.

Figure 4.8. CD spectra of mCM variants (15 µM) in presence versus absence of TSA (30 µM).

Figure 4.9. Chemical denaturation of Wt mCM and its mutants (10 µM) monitored by UVCD at 222 nm wavelength with increasing concentrations of GdnHCl (0-6 M). Solid fitted lines represent a dose-response curve of non-linear regression analysis using OriginPro 8 software.

Figure 4.10. Thermal denaturation of Wt mCM. (*Left*) Far UV CD spectra at temperature intervals during the heating cycle (25-90 ^{0}C, solid coloured lines) and the reverse cycle (90-25 ^{0}C, dashed coloured lines) measured in 10 mM phosphate and NaF buffer, pH 7.2, concentration 10 µM. (*Right*) Thermal melting curve of Wt mCM following changes in the ellipticity at 222 nm wavelength. Ellipticity was monitored continuously in the temperature range from 25 to 90 ^{0}C (●) and during a reverse scan from 90 to 25 ^{0}C (○).

Figure 4.11. Thermal denaturation of Mutant1. (*Left*) Far UV CD spectra at temperature intervals during the heating cycle (25-90 ^{0}C, solid coloured lines) and the reverse cycle (90-25 ^{0}C, dashed coloured lines) measured in 10 mM phosphate and NaF buffer, pH 7.2, concentration 10 µM. (*Right*) Thermal melting curve of Wt mCM following changes in the ellipticity at 222 nm wavelength. Ellipticity was monitored continuously in the temperature range from 25 to 90 ^{0}C (●) and during a reverse scan from 90 to 25 ^{0}C (○).

Figure 4.12. Thermal denaturation of Mutant2. (*Left*) Far UV CD spectra at temperature intervals during the heating cycle (25-90 ^{0}C, solid coloured lines) and the reverse cycle (90-25 ^{0}C, dashed coloured lines) measured in 10 mM phosphate and NaF buffer, pH 7.2, concentration 10 µM. (*Right*) Thermal melting curve of Wt mCM following changes in the ellipticity at 222 nm wavelength. Ellipticity was monitored continuously in the temperature range from 25 to 90 ^{0}C (●) and during

List of tables

List of abbreviations

AA	Amino acid
ANS	8-Anilinonaphthalene-1-sulfonic acid
BSA	Bovine serum albumin
CD	Circular dichroism
Cys	Cysteine
DNA	Deoxyribonucleic acid
DMF	Dimethylformamide
dsDNA	double-stranded deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ER	Electromagnetic radiation
FOP	Fully ordered proteins
FRET	Förster resonance energy transfer
FWHM	Full-width half-maximum
Gdn.HCl	Guanidine hydrochloride
IDP	Intrinsically disordered proteins
IDR	Intrinsically disordered regions
IPTG	Isopropyl-ß-D-1-thiogalactopyranoside
IRF	Instrument response function
mCM	monomeric chorismate mutase
MD	Molecular dynamics
MALDI	Matrix Assisted Laser Desorption/Ionization
mMjCM	monomeric chorismate mutase
MEM	Maximum entropy method
Ni-NTA	Nickel- Nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NLLS	Non-linear least square
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline

PCR	Polymerase chain reaction
ProCharTS	Protein charge transfer spectra
PDB	Protein data bank
RET	Resonance energy transfer
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SDM	site-directed mutagenesis
ssDNA	Single-stranded deoxyribonucleic acid
TCSPC	Time-correlated single photon counting
Trp	Tryptophan
TSA	Transition state analogue
UV	Ultraviolet
Wt	Wild type

Statement regarding new finding

1. The tryptophan residues in mCM is solvent accessible and flexible.
2. The two tryptophans in wild type mCM undergo homoFRET.
3. Ligand binding doesn't directly affect the tryptophan microenvironment.
4. Time-resolved anisotropy measurements reveal ligand-induced global compaction in mCM as a mechanism for disorder-to-order transition for effective catalysis.
5. The ability to undergo ligand-induced global compaction and thereby disorder-to-order transition is drastically lowered in rationally designed mutants (W24K/C69A; W24K/C69A/A6C).
6. The lowering of ligand-induced global compaction is correlated with decrease in catalytic function of the enzyme.
7. The mutants also show changes in secondary structure, chemical stability and deviation in hydrophobic core geometry, which in turn could affect enzyme activity.
8. The chemical denaturation states of Wt mCM and mutants are different from thermal denaturation. All protein species show cooperative two-state chemical denaturation while non-cooperative linear thermal denaturation.
9. mCM Mutant3 (W24K/C69A/A6C) show some degree of compensation in perturbation of alpha helicity, chemical stability and enzyme activity when compared to mCM Mutant2 (W24K/C69A).
10. The mCM Mutant2 (W24K/C69A) displayed greatest loss in alpha helicity and showed least chemical stability among all mutants.
11. Few crucial non-active site residues impact conformational integrity and catalytic function.
12. Alpha helicity, segmental dynamics, disorder-to-order transition and ligand induced global compaction are essential determinants for effective catalysis in mCM.

ABSTRACT

Intrinsically disordered proteins (IDPs) are a class of proteins gaining immense attention over the last two decades due to its unique structural characteristics and relevance in many major human diseases. They are known to carry out a variety of crucial biological functions viz. signalling, regulation and transcription in spite of lacking a well-defined 3-dimensional structure. An IDP exists in a flexible disordered state in its native conformation and can adopt ordered structure(s) after binding to a suitable partner biomolecule (protein/DNA/RNA) or ligand. Nevertheless, they have been observed to perform catalytic activities in very rare cases. Monomeric chorismate mutase (mCM) is one such rare disordered enzyme that efficiently catalyses the conversion of chorismate to prephenate, a crucial step in the biosynthesis of aromatic amino acids. It possesses characteristic of a molten globule (classified under IDPs), which are generally found as intermediate structures in protein folding. Interestingly, it retains a relatively flexible structure even after complex formation with its substrate/ligand. It is natural for intrinsically disordered enzymes to draw considerable interest and attention in the field.

mCM is obtained through an engineered process of genetic selection in combination with structure-based design from the enzyme AroQ chorismate mutase (*Methanococcus jannaschii*), known as MjCM. This approach converts the dimeric helical bundle enzyme (MjCM) to the monomeric four helix bundle protein (mMjCM or simply mCM) while still retaining its enzyme activity. The simple four helical bundle structure and the presence of only two tryptophan and a single cysteine residue allowed us to investigate any environment and ligand associated structural transitions in mCM through fluorescence methods viz. time-resolved FRET and anisotropy. The dominant view in enzymology being that an increase in structural disorderedness tends to decrease catalytic efficiency; therefore, in the case of the mCM with properties of a molten globule, what then contributes to the enzyme catalysis? **This is the broad theme of my thesis**. Does conformational plasticity play a lead role in the catalytic mechanism? If yes, what kind of conformational diversity drives catalysis? Global vs. local conformational plasticity- Which holds the key to regulation of enzyme activity? How is the flexibility affected in the presence of ligands? The answer of some of these challenging questions in the field is far from being understood and is a matter of intense

debate. With regard to this, the project has been divided into two interconnected working modules, each addressing specific aims:

Working module 1: What is the nature of structural plasticity mechanisms in engineered monomeric chorismate mutase (mCM) that enable catalysis?

Working module 2: How do designed mutations alter catalysis? How are the segmental dynamics, conformational adaptation, and disorder-to-order transition correlated with changes in catalysis?

Capturing the structural transitions in mCM associated with its catalysis can present us with a much better understanding of how disordered enzymes function efficiently in spite of such flexibility.

In this thesis, we probed the structure and dynamics of mCM using the intrinsic fluorophore tryptophan and the extrinsic fluorophore dansyl (labeled to cysteine residue), which also served as a FRET pair in our experiments. Through site-directed mutagenesis, single tryptophan and single cysteine mutant pair of mCM protein were generated at specific locations (based on NMR structure with substrate analogue) such that FRET between a Trp (donor)-Dansyl (acceptor) pair can be measured effectively and without perturbation of the substrate binding site of the enzyme. FRET acts as a molecular ruler and measures ligand associated structural transitions in mCM in terms of linear distance change. Three mutants namely Mutant1 (W24K), Mutant2 (W24K/C69A) and Mutant3 (W24K/C69A/A6C) were extensively studied and their structural and functional properties compared with the parent wild-type mCM enzyme (Wt mCM).

Steady-state fluorescence measurements showed distinct variations in fluorescence intensity and emission maxima in the mCM variants indicating subtle changes in the local environment around the tryptophan residue(s) after mutations. The lower fluorescence intensity of the two closely placed (~8 Å) tryptophan residues in the Wt mCM compared to the single tryptophan residue in the mutants suggests the presence of Trp-Trp homo-FRET (self-quenching) in Wt mCM. This was confirmed by measuring the 310/295 steady state anisotropy ratio for Wt mCM tryptophans which provided a value greater than 2 (2.42), expected from systems exhibiting homo-FRET.

Fluorescence assays with the hydrophobic probe ANS showed noticeable difference in fluorescence intensity of the different mutants when compared to the Wt mCM suggesting change in accessibility of ANS to hydrophobic patches in mCM probably due to conformational changes in the mutants. Interestingly, substantial decrease in ANS fluorescence was observed in Wt mCM as well as the mutants after a Transition-State Analogue (TSA) was first allowed to bind mCM variants before ANS binding. This suggests overlapping binding site for TSA and ANS in mCM, basically the hydrophobic core of the enzyme. Since ANS fluorescence intensity in presence of TSA is drastically reduced in all the mutants, they seem to have retained their ability to bind the ligand even after secondary structure and conformational changes. Significant changes in secondary structure of mCM was also observed in Mutant2 and Mutant3 through CD measurements. A notheworthy decrease in α-helicity was observed in the mutants with cysteine mutants, Mutant2 (W24K/C69A) and Mutant3 (W24K/C69A/A6C). Interestingly, some degree of α-helicity was recovered in the Mutant3 where a single cysteine residue is introduced at position 6.

Time-Resolved Fluorescence measurements could sensitively detect local changes in structure of mCM after TSA binding. The single tryptophan mutant (W24K) of mCM displayed a substantial shift between the Trp-Cys FRET pair of approximately 6 Å after addition of the ligand. Moreover, the triple mutant (Mutant3) also showcased a decrease in FRET distance from 29.4 Å to 25.2 Å after large excess of the ligand was used for binding. Time-resolved fluorescence anisotropy decay measurements revealed global structural alternations in mCM after ligand (TSA) binding in the Wt mCM as well as the mutants. A long correlation time (θ) of 8.6 ns, corresponding to the global rotation of mCM, decreased to 6.3 ns after ligand binding in Wt mCM. This decrease in the long component of correlation time indicates the formation of a more compact global conformation. A similar pattern of decrease in the longer correlation time (θ) was observed in the mutants after ligand binding, but they reported larger values for the longer correlation times (θ) before (10-11 ns) as well as after (8.5-10.5 ns) ligand binding in comparison to Wt mCM. Thus, it seems that the mutants possess a less compact structure than the Wt mCM even after TSA binding.

Our enzyme kinetics studies shed light on the functional properties of Wt mCM and its mutants. Wt mCM reported K_m and k_{cat} values of 195 μM and 8.8 s^{-1} respectively, close to previously reported values by the Hilvert group (ETH Zurich). The enzyme efficiency (k_{cat}/K_m) was calculated to be 45128.3 $M^{-1}s^{-1}$. Mutant1 retained 85% of the enzyme activity, whereas drastic drop in enzyme efficiency to ~2.5% in Mutant2 and ~11% in Mutant3 was observed. These experimental results also emphasize the role of the single cysteine residue (Cys 69) in maintaining the structural and functional integrity of mCM, even though Cys 69 doesn't form a part of the enzyme active site.

The overarching theme that has emerged from the series of carefully designed experiments in my thesis is that the integrity of the secondary structure and ligand-induced global compaction (disorder-to-order transition) play a crucial role in catalysis and that even minor perturbation in regions distant from the active site can significantly impact the enzyme efficiency by changing the conformation, segmental dynamics, ligand-induced global compaction and stability.

Thesis Objectives

The work presented herein is in partial requirement for the completion of PhD thesis from the Department of Biosciences and Bioengineering, Indian Institute of Technology, Guwahati, India under the guidance of Prof R. Swaminathan. The current thesis work describes the structure, function, and dynamics of an engineered chorismate mutase enzyme with the properties of a molten globule.

Proteins are intrinsically flexible molecules spanning a continuum of partially or totally disordered to finely folded, native states. Enzymes are predominantly associated with species possessing a well-defined structure due to its requirement of precise positioning of active site residues for effective catalysis. The engineered monomeric chorismate mutase (mCM) displays the properties of a molten globule, but very efficiently catalyzes the rearrangement of chorismate to prephenate similar to the wild type dimer. mCM has therefore emerged as a valuable model system to study enzymatic catalysis by an intrinsically disordered protein (IDP). The IDPs represent complex systems with exceptional spatio-temporal heterogeneity that are chaotic and are tunable by various means.

The thesis investigates the effects of **conformational diversity** and **rationally designed mutations** on **enzyme function**. The project has been divided into **two interconnected working modules**, each addressing specific aims:

Working module 1: What is the nature of structural plasticity mechanisms in engineered monomeric chorismate mutase (mCM) that enable catalysis?

Working module 2: How do designed mutations alter catalysis? How are the segmental dynamics, conformational adaptation, and disorder-to-order transition correlated with changes in catalysis?

The disentanglement of IDP related complexity in modulating enzyme catalysis is the central thrust of the thesis analyzed by highly sensitive fluorescence spectroscopy coupled with site-directed mutagenesis.

CHAPTER 1

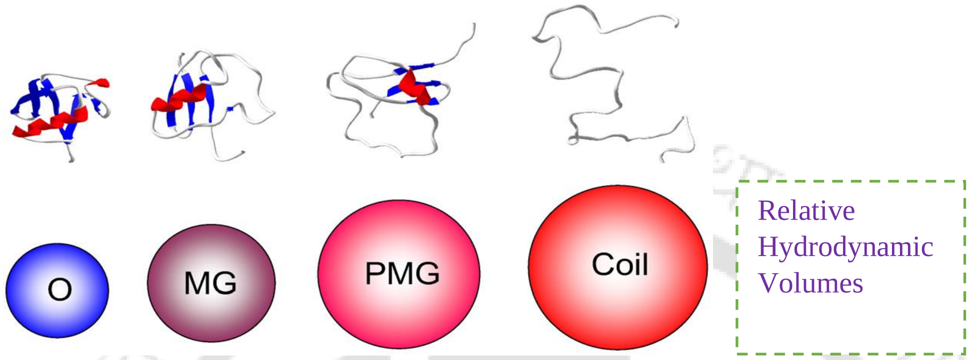
Introduction
and
Review of Literature

1.1. INTRINSICALLY DISORDERED PROTEINS

Over the past century, the function of a protein was defined by its ability to adopt and adhere to a characteristic 3-dimentional (3D) structure that subsequently drove its function. This phenomenon termed as the "structure-function paradigm" was the cornerstone on which our understanding of protein function was laid upon. Fluidity in structure was generally regarded as an anomaly and therefore ignored or neglected. In 1950, a study done on serum albumin by Karush, reported that albumin possessed high-affinity binding capacity for small, hydrophobic, particularly anionic molecules in an almost universal manner. Interestingly, molecules with very different shapes demonstrated competitive binding. Through his experimental observations, Karush arrived at the conclusion that albumin's binding site can assume a variety of different configurations of similar energies in equilibrium with each other. When albumin encounters a given hydrophobic anion, the most fitting configuration is selected from the protein's structural ensemble. This phenomenon was referred to as configurational adaptability by Karush (Karush 1950). In an independent work by Koshland, he also proposed the concept of configurational adaptability, but referred to this process as induced fit (D. E. Koshland 1958). Both configurational adaptability and induced fit are the first known models that suggested the requirement of significant conformational change for protein function to occur. Subsequent work on other proteins lead researchers to believe that function depends on prior 3D structure, or on structural accommodations within a prior 3D structure, or on regulatory shifts between alternate structures (Dunker *et al.*, 2001).

With the advent of the new millennium, the explosion in information on proteins that deviate from the universal structure-function paradigm urged the scientific community to expand their viewpoint on the relationship between protein structure and its function. Many examples where non-structured segments of proteins or proteins that are completely unstructured, gave rise to protein function emerged.

Early research work built a backbone around these non-structured proteins and sparked a general interest and appreciation for them. These atypical proteins came to be collectively known as intrinsically disordered proteins (IDPs, originally termed, unstructured). It is now accepted that to be "intrinsic disordered", a protein (or protein region) has to be biologically active, yet exists as collapsed or extended, dynamically mobile conformational ensemble, either at the level of secondary or tertiary structure (V. N. Uversky, Gillespie, and Fink 2000)(Daughdrill et al. 2005)(Wright and Dyson 1999)(Tompa 2002)(A K Dunker et al. 1998)(Dunker *et al.*, 2001)(V. Uversky and Keith Dunker 2010).



Source: Biochim et Biophys Acta | 2010 | Vol. 1804 | (V. Uversky and Keith Dunker 2010)

Figure 1.1. Illustrative examples of Intrinsically Disordered Proteins (IDPs). Collapsed disorder (molten globule like, MG), extended disorder (pre-molten globule-like, PMG) and statistical coil (coil) disorder. For comparison, an ordered globular protein (O) of same length is also shown.

IDPs and intrinsically disordered regions (IDRs) exist as dynamic ensembles, resembling “protein clouds” (A Keith Dunker and Uversky 2010), in which the atom positions and backbone Ramachandran angles vary significantly over time with no specific equilibrium values and undergo non-cooperative conformational changes. Although highly dynamic, often IDP structures can be described well by a limited number of lower-energy conformations (Choy and Forman-Kay 2001)(Huang and Stultz 2008). A random coil refers to a long-chain polymer wherein the entire backbone lacks a well-defined structure, which is derived from the polymer theory. This concept assumes that the polymer displays no significant intra-protein, non-local interactions and exists as a freely-jointed, freely-rotating chain in a "theta solvent" undergoing unrestricted sampling of the sterically allowed regions of the Ramachandran plot. In Flory's model of the polymeric random coil, each monomer subunit is randomly oriented with respect to the neighboring monomer subunits known as the "isolated pair hypothesis." The random coil model is bereft of the idea that conformational propensities may be modulated by the interaction of peptide/protein with the solvent and with nearest-neighbors. It is worthwhile to consider that unfolded proteins have greater structural diversity than assumed previously. A compelling case in this regard has been made in an elegant review by Toal and Scheitzer-Stenner where they discuss the experimental, computational, and bioinformatic evidence for the necessity of avoiding the term random coil for unfolded proteins/peptides at the local level (Toal and Schweitzer-Stenner 2014).

The Scheraga's 'statistical coil' model would imply that Levinthal-like number of denatured conformations will exist (Pappu, Srinivasan et al. 2000, Zaman, Shen et al. 2003), indicating significant conformational flexibility. However, correlations between neighboring residues would influence this number. Therefore, assessing the influence of nearest neighbors on the conformational ensemble in intrinsically disordered proteins is crucial for a thorough understanding of the statistical coil state of the IDPs. conformational sampling of individual residues in a disordered state is much more restricted than the random coil model. Therefore, the 'statistical coil' depicts all aspects of the disordered state. As a result, extended regions (statistical coil-like) as well as partially collapsed regions with residual secondary structure (pre-molten globule like), domains with noticeable secondary structure and compaction (partially folded or molten globule-like), and domains with poorly packed side chains are included in the current view of intrinsic disorder as depicted in **Figure 1.1** (V. N. Uversky, Gillespie, and Fink 2000)(Uversky and Dunker, 2010)(V. N. Uversky 2002a)(V. N. Uversky 2002b)(A. Keith Dunker et al. 2008)(A Keith Dunker and Obradovic 2001).

In the presence of a binding partner, IDPs can undergoes a disorder-to-order transition (Daughdrill et al. 2005)(Wright and Dyson 1999)(Dunker *et al.*, 2001)(Brown et al. 2002)(Dunker, Brown and Obradovic, 2002)(A Keith Dunker et al. 2002)(V. Uversky, Oldfield, and Keith Dunker 2005)(A Keith Dunker et al. 2005)(Uversky, Oldfield and Dunker, 2005)(Iakoucheva et al. 2002)(A Keith Dunker and Obradovic 2001)(Spolar and Record 1994)(Pontius 1993)(Dyson and Wright 2002)(Dyson and Wright 2005)(Plaxco, Gross, and Michael 1997). However, instead of binding associated folding, some IDPs form highly dynamic complexes with their partners (Nash et al. 2001)(Baker et al. 2007)(Borg et al. 2007)(Sigalov 2010) (Mittag et al. 2008)(Mittag, Kay, and Forman-Kay 2010)(Mittag et al. 2010).

1.1.1.Experimental methods used to characterize IDPs

X-ray crystallography: More than 20 years ago, during protein crystal structure determination, some protein segments were discovered to yield no discernible electron density and yet participate in protein function. A common reason for missing electron density is that the unobserved atom, side chain, residue, or region fails to scatter X-rays coherently due to variation in position from one protein to the next, i.e., the unobserved atoms are disordered. Although some of these missing density regions may correspond to wobbly, structured domains rather than to intrinsically disordered ensembles, their presence is quite rare among the long regions of missing electron density (Radivojac et al. 2004). Data acquired from X-ray diffraction needs further validation from other experimental techniques to ascertain whether a region of missing electron density is intrinsically disordered or is the result of some technical difficulties (V. N. Uversky 2011).

NMR spectroscopy: One of the first evidence of functional disorder by NMR was found in 1978, the same year that functional disorder was suggested by X-ray crystallography. It was revealed by NMR that the functional and highly charged tail of Histone H5 protein was disordered (Aviles *et al.*, 1978). From that time onwards, NMR 3D structural determination has led to the characterization of several proteins containing functional disordered regions (Muchmore et al. 1996)(Riek et al. 1996). By utilizing a variety of different isotopic labelling and pulse sequence experiments, NMR can provide motional information of proteins on a residue-by-residue basis (Ishima and Torchia 2000). A particularly useful measurement is ^{15}N-^{1}H hetereonuclear NOE, which generates positive values for the slowly tumbling ordered residues and negative values for rapidly tumbling disordered residues (Ishima and Torchia 2000)(Bracken 2001). For an entire protein, NOE data is therefore evident as a series of consecutive positive values for the ordered segments and consecutive negative ones for the disordered segments (Muchmore et al. 1996). One relevant point for us to consider is that molten globules (refer section 1.2.) represent a tough terrain when structural elucidation of these disordered proteins are undertaken by NMR, while statistical-coil disorder can be studied without much difficulty. This is because molten globules are heterogeneous with structural interconversions on the millisecond timescale, which leads to extreme broadening of the side-chain NMR peaks. Moreover, an opposite effect of a lack of chemical shift dispersion (rather than extreme line broadening) was observed for NMR backbone data on molten globules, further adding to this challenge (Dunker *et al.*, 2001). Nevertheless, NMR is overall a more reliable technique for characterization of disorder than X-ray diffraction. The affirmation that disorder existed, through NMR experiments, apart from the evidence obtained from X-ray diffraction, made a major impact in the growth of literature in IDPs (Plaxco, Gross, and Michael 1997)(Wright and Dyson 1999).

Circular dichroism (CD) spectroscopy: Disordered proteins can be distinguished from ordered ones by structural information from CD spectra. Near-UV CD presents sharp peaks for aromatic groups in ordered proteins, but as a result of motional averaging, these peaks are not observed for molten globules and statistical coils. Whereas, far-UV CD spectra can distinguish ordered and molten globular forms from statistical coil by estimates of their secondary structure (Kuwajima 1977) (Dolgikh et al. 1981)(Ohgushi and Wada 1983). However, this technique lacks residue-specific information and hence cannot provide definitive information for proteins that contain both ordered and disordered regions.

Protease digestion: Information on protein structure and its flexibility have been derived from protease digestion since the early 20th century (Markus 1965). Further studies by Fontana et al. strongly suggests that flexibility, not mere surface exposure, is the determining factor for digestion of possible cut sites (Fontana, de Laureto, and De Filippis 1993)(Fontana et al. 1997)(Hubbard, Eisenmenger, and Thornton 1994)(Hubbard, Beynon, and Thornton 1998). They demonstrated that when the F helix of myoglobin is converted to a disordered state in apomyoglobin, a significant increase in digestion rates occurs. This disordered region arising from the F helix presented cut loci for several different proteases. Hence, they concluded that hypersensitivity to proteases is a reliable parameter to assess the presence of disorder in proteins. Moreover, protease digestion also provides position-specific information. One limiting factor to the extensive use of this technique for disorder determination is the requirement for protease-sensitive residues within a protein. A combination of protein digestion with techniques like X-ray crystallography, CD spectroscopy and mass spectrometry can provide more useful insights into protein disorder (Dunker *et al.*, 2001).

Stoke's radius determination: For a given molecular weight, an abnormally large Stoke's radius indicates protein disorder. Statistical-coil disorder can be detected by measurement of the Stoke's radius by methods viz. size exclusion chromatography and small-angle X-ray scattering (Schweers et al. 1994). Independent evidence of statistical coil structure has been obtained by a combination of the above-mentioned methods along with CD spectroscopy (Kriwacki et al. 1997)(Weinreb et al. 1996).

Computational prediction tools: In the subsequent years following the discovery of IDPs, a number of computational methods have also been developed that exploit the sequence signatures of disorder to predict whether a protein is disordered, given its amino acid sequence. The table below (**Table 1**) highlights a few popular methods and their special predicting features.

Table 1. Computational predictors for structural disorder. (Adapted from: Ferron, Longhi, Canard, & Karlin, 2006)

Predicter	What is predicted	Basis of prediction
PONDR	All regions that are not rigid including statistical coils, partially unstructured regions, and molten globules	Local aa composition, flexibility, hydropathy, etc.
IUPred	Regions that lack a well-defined 3D-structure under native conditions	Energy resulting from inter-residue interactions, estimated from local amino acid composition
DisEMBL	LOOPS (regions devoid of regular secondary structure); HOT LOOPS (highly mobile loops); REMARK465 (regions lacking electron density in crystal structure)	Neural networks trained on X-ray structure data
s2D	Predict secondary structure and intrinsic disorder in one unified statistical framework based on the analysis of NMR chemical shifts	Neural networks trained on NMR solution-based data
FoldIndex	Regions that have a low hydrophobicity and high net charge (either loops or unstructured regions)	Charge/hydropathy analysed locally using a sliding window

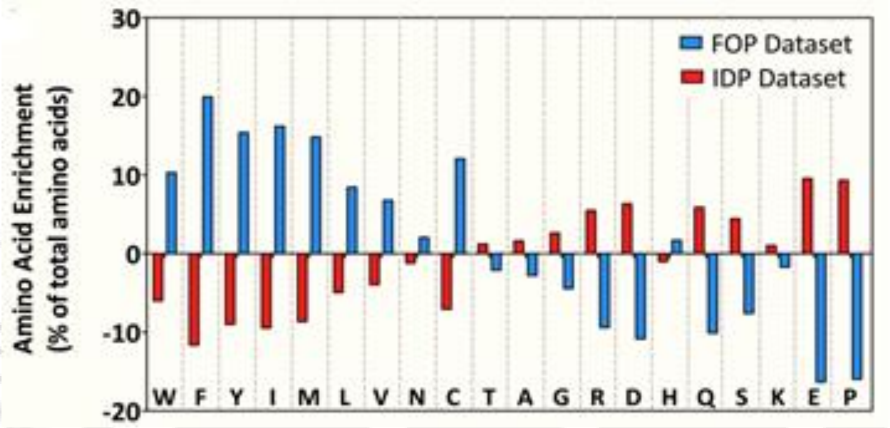
Early use of these disorder predictors could reveal the abundance of disordered proteins in the proteome. An analysis of the SwissProt database yielded more than 15,000 proteins with regions of disorder of at least 40 consecutive amino acids, with about a

1000 proteins displaying very high disorder scores (P Romero et al. 1998). Moreover, intrinsic disorder appeared to be a common phenomenon at the genome level, with a higher percentage of disordered regions observed in eukaryotes (27–41%) compared to archaea (4–24%) or bacteria (2–21%) (A K Dunker et al. 2000). Further bioinformatics studies suggested that more than half of eukaryotic proteins possess long disordered regions and about 25-30% of eukaryotic proteins were highly disordered (A K Dunker et al. 2000)(Oldfield, Cheng, Cortese, Brown, et al. 2005)(Oldfield, Cheng, Cortese, Romero, et al. 2005). Signalling proteins were found to harbour long disordered regions in about 70% of the cases (Iakoucheva et al. 2002). One related speculation is that more disorder is needed for signalling and coordination among the various organelles in the more complex eukaryotic domain (A Keith Dunker and Obradovic 2001). Although IDPs don't exhibit a fixed tertiary structure under physiological conditions, its high natural abundance is a clear indication of their importance in biological functions (V. N. Uversky, Gillespie, and Fink 2000)(Wright and Dyson 1999)(Tompa 2002)(A. Keith Dunker et al. 2001)(Brown et al. 2002)(A Keith Dunker et al. 2002)(A Keith Dunker et al. 2005)(V. N. Uversky 2002a)(V. N. Uversky 2002b)(Iakoucheva et al. 2002)(Dyson and Wright 2002)(Dyson and Wright 2005)(Fink 2005).

1.1.2.Characteristic features of IDPs

Charge and hydrophobicity: A striking feature of highly disordered proteins is their high content of charged residues (V. N. Uversky, Gillespie, and Fink 2000). In fact, charged residues are one of the defining attributes of IDP structure and function. The high net charge is crucial for extended conformation to occur (Mao et al. 2010). In aqueous solutions, a number of cases have been reported that show formation of heterogeneous ensembles of collapsed structures from protein sequences that are rich in uncharged, polar amino acids and devoid of canonical hydrophobic residues. (Crick et al. 2006)(Wang et al. 2006)(Möglich, Joder, and Kiefhaber 2006)(Vitalis, Wang, and Pappu 2007)(Mukhopadhyay et al. 2007)(Dougan et al. 2009)(Walters and Murphy 2009). Moreover, recent studies have concluded that charge content can modulate the intrinsic preference of polypeptide backbones for collapsed structures (Mao et al. 2010). A good example validating this conclusion comes from proteins known as Nucleoporins (Nups). They contain large intrinsically disordered domains with multiple Phe-Gly repeats (FG domains). Highly charged Nups displayed more dynamic, extended coil conformations whereas the Nups with low charge contents gave rise to globular configurations. Distinct categories of intrinsically disordered structures were adopted by Nups depending on the content of charged residues (Yamada et al. 2010).

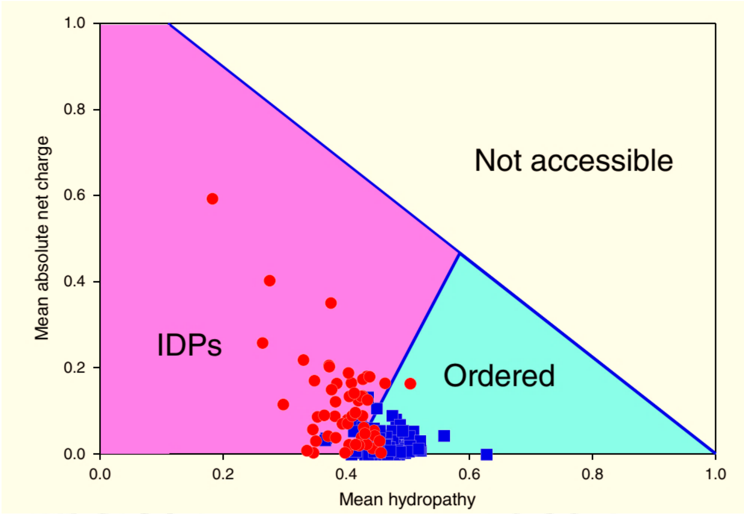
Amino acid sequence can also determine the tendency of a protein to fold or not to fold. In disordered proteins, amino acids that promote order such as Ile, Leu, Val, Trp, Tyr, Phe, Cys, and Asn are significantly depleted and those known to be associated with disorder such as Ala, Arg, Gly, Gln, Ser, Glu, Lys, and Pro are significantly enriched as shown in **Figure 1.2** (Dunker *et al.*, 2001)(R. M. Williams et al. 2001)(Pedro Romero et al. 2001)(Radivojac et al. 2007)(Dunker *et al.*, 2008)(Vacic et al. 2007).



Adapted from: Plos one | 2016 | vol. 11 | (Bürgi et al. 2016)

Figure 1.2. Amino acid enrichment in IDPs showing the relative amino-acids composition of IDPs. The enrichment is calculated by the formula: 100 - (% aa in IDP*100/% aa in total dataset). FOP stands for fully ordered protein.

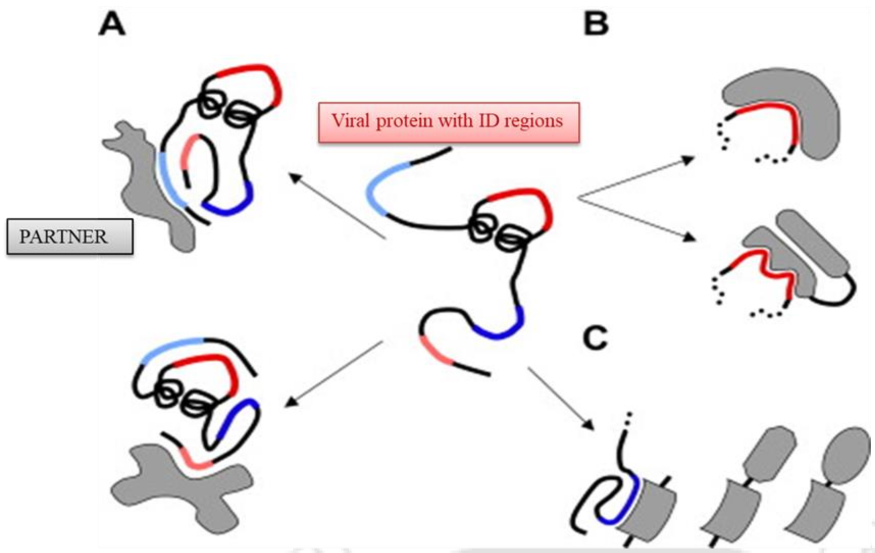
Analysis of 275 natively folded and 91 natively unfolded proteins by Uversky et al. concluded that a combination of low mean hydrophobicity (leading to low driving force for protein compaction) and high net charge (leading to strong electrostatic repulsion) is an important prerequisite for the absence of compact structure in a protein as depicted in **Figure 1.3** (V. N. Uversky, Gillespie, and Fink 2000)(Uversky et al., 2000).



Adapted from: Biochimica et Biophysica Acta | 2013 | Vol. 1834 | (V. N. Uversky 2013)

Figure 1.3. Charge-hydropathy plot for ordered proteins (blue squares) and IDPs (red circles). The area accessible to sequences encoding ordered and disordered proteins are shown in contrast colours, cyan and pink respectively, is shown as a light cyan triangle, whereas the area accessible to sequences encoding IDPs is depicted as light pink pentagon. These two areas are defined by two boundaries, the known boundary separating ordered proteins and IDPs and the boundary showing logical limits of the CH-space.

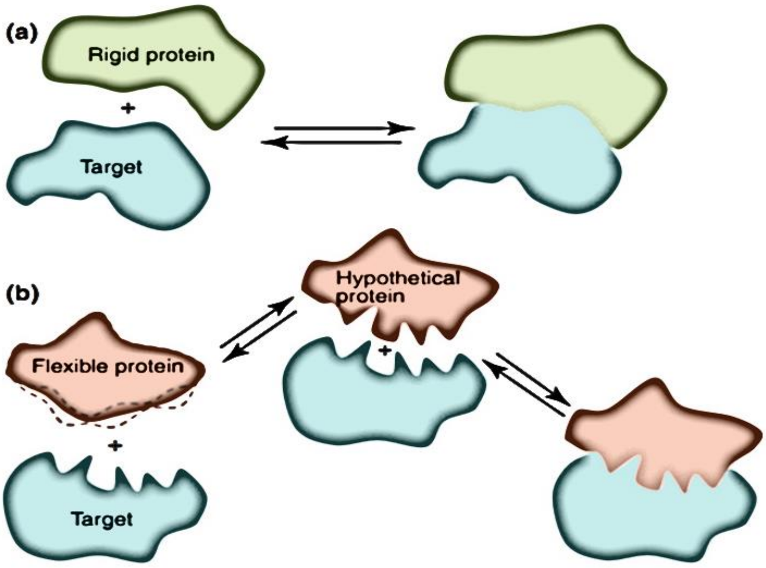
Binding promiscuity: The inherit flexible nature of IDPs allows them to easily interact with multiple partners, behaving promiscuously. This is a huge advantage as many IDPs serve in protein interaction and signalling networks as nodes or hubs (A Keith Dunker et al. 2005)(Uversky, Oldfield and Dunker, 2005)(Patil and Nakamura 2006)(Ekman et al. 2006)(Haynes et al. 2006)(Dosztányi et al. 2006)(Singh, Ganapathi, and Dash 2007)(Singh and Dash 2007). p53, p21, p27, α-synuclein, BRCA1, estrogen receptor are some of the known disordered hub proteins with a large repertoire of binding partners. The promiscuity of IDPs allow them to either act as a hub to bind many partners or multiple IDPs can bind to a single structured hub protein as depicted in **Figure 1.4**. IDRs (Intrinsically Disordered Regions within a protein) can also serve as linkers between ordered functional domains allowing unhindered movement of the domains with respect to each other as well as increase the binding diversity of the protein (A Keith Dunker et al. 2005)(V. Uversky, Oldfield, and Keith Dunker 2005).



Adapted from: Microbes and Infection | 2010 | Vol.12 | (Vidalain and Tangy 2010)

Figure 1.4. Viral proteins to bind many partners through intrinsically disordered regions (IDRs). Disordered regions can adopt different fold upon binding to different partners (A), small linear motifs within disordered regions can adopt different fold to interact with different partners (B), viral proteins can interact with multiple cellular proteins carrying a conserved domain through these small linear motifs (C).

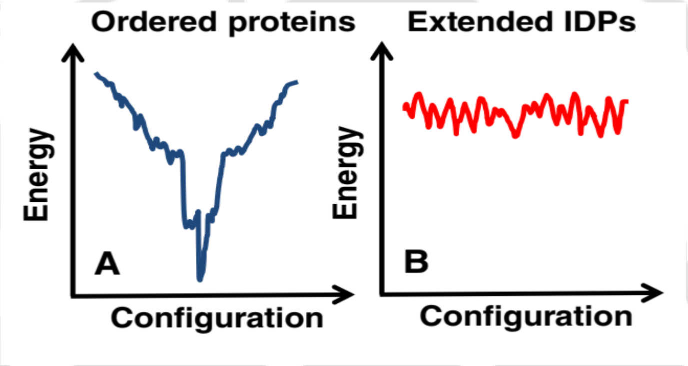
Kinetic advantage of intrinsic disorder: The structural flexibility and ability of IDPs to undergo disorder-to order transitions provide them with unique functional advantages. Literature on IDP functional features by several groups present the following advantages: (i) decoupled specificity and strength of binding sets ground for the high-specificity low-affinity interactions; (ii) increased capture radius for a specific binding site as compared to ordered protein with its restricted conformational freedom; (iii) a greater capture radius and the ability to spatially search through interaction space allows increased speed of interaction; (iii) increased interaction (surface) area per residue; (iv) the ability for one-to-many and many-to-one interactions; (vi) fast binding kinetics and so on as depicted in **Figure 1.5** (Wright and Dyson, 1999)(Dunker *et al.*, 1998)(Dunker *et al.*, 2001) (Uversky and Dunker, 2010)(Romero *et al.*, 2001)(Cortese *et al.*, 2008). The fast binding kinetics and increased capture radius is responsible for the "**fly-casting mechanism**" of protein binding. According to this mechanism, the unfolded polypeptide first binds weakly at a relatively large distance from the actual binding site and then gradually folds as the protein approaches the binding site (Shoemaker *et al.*,2000).



Adapted from: Trends in Biochem Sciences │Feb 2012 │Vol. 37 (Zhou 2012)

Figure 1.5. Binding affinity and specificity of a rigid ordered protein and a flexible disordered protein. (a) Interaction with target done via a relatively smooth interaction surface by a rigid protein (b) Left: Interaction with target after sampling multiple conformations (especially around binding site) by a flexible protein. The dashed curves represent alternative conformations (as illustrated by dashed curves representing alternative conformations. Right: After initial binding, flexibility aids the protein in wrapping around the protrusions of the target, which then gives rise to a convoluted interaction surface and high specificity. Middle: a hypothetical protein (not flexible) that adopts the bound conformation even when unbound would have a much higher free energy than a flexible protein. Unlike the flexible protein, this hypothetical protein would have a high binding affinity for the target, yet comparable specificity.

Energy Landscape: Proteins can sample a large number of conformations in their lifespan and the presence of any specific conformation is determined by the underlying energy landscape (Fisher and Stultz 2011). The dependence of the free energy on all the coordinates determining protein conformation is described by this landscape (V. N. Uversky 2013). In well-folded proteins, the number of conformational states accessible by the polypeptide chain is reduced as the protein approaches the native state; therefore well-folded ordered proteins can be described by a funnel-like energy landscape with a well-defined global energy minimum as shown in **Figure 1.6** (Radford 2000)(Jahn and Radford 2005). On the contrary, the free energy corresponding to the dynamic ensemble of the large number of conformations characteristic of IDPs can be described as a large "hilly plateau", where the forbidden conformations are represented by the hills on the plateau (V. N. Uversky, Oldfield, and Dunker 2008)(Fisher and Stultz 2011)(Turoverov, Kuznetsova, and Uversky 2010). The energy landscape of an IDP is characterized by numerous local energy minima and hence is relatively flat. The multiple local energy minima create a highly frustrated system without any stable well-folded conformation (V. N. Uversky 2013). This gives rise to an exceptionally environment sensitive energy landscape that in turn determines the conformational plasticity of IDPs. Different environmental factors can produce different effects on the energy landscape, causing the depth/height of the local energy minima to vary accordingly. Hence, IDPs can adopt different folds depending on the environmental conditions and also upon interaction with different binding partners. It is assumed that a particular binding interaction can affect the folding landscape in a unique way, promoting the formation of a specific structure in a template dependent manner (V. N. Uversky 2013).



Adapted from: Biochimica et Biophysica Acta | 2013 | Vol. 1834 | (Uversky, 2013)

Figure 1.6. Energy landscape of ordered proteins and IDPs. The energy versus configuration plots show the folding energy landscapes of a typical globular ordered protein (A) and of a typical natively unfolded protein (B).

Abnormal electrophoretic mobility and heat resistance: Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) is a well-established method routinely used to assess molecular weight (M_w) of proteins. Interestingly, owing to their highly charged amino acid composition, IDPs bind relatively less SDS than normal globular proteins. Experiments reveal that their apparent M_w is often 1.2–1.8 times higher than the actual one measured by mass spectrometry or determined theoretically from sequence data (Tompa 2002). Abnormal electrophoretic mobility is a good indication of the presence of substantial disorder in a protein (Sieber et al. 2011)(Popelka, Uversky, and Klionsky 2014).

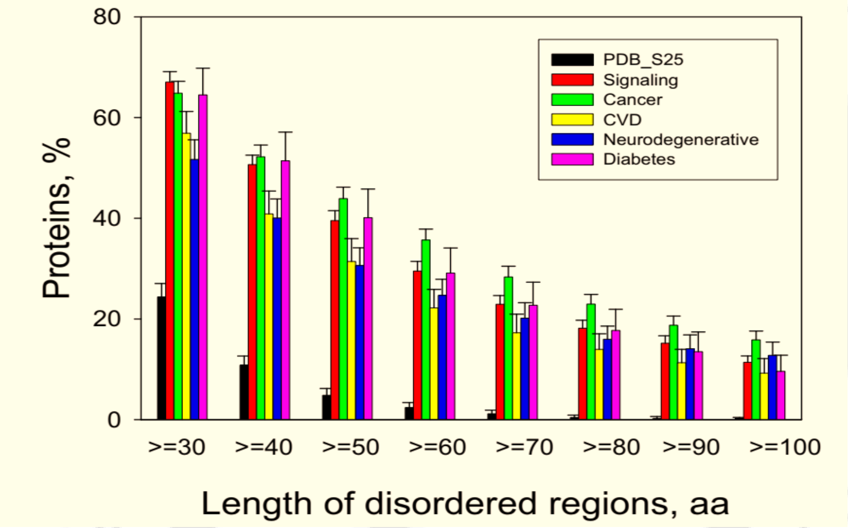
Several previous studies indicated that IDPs are stable to heat denaturation. For example, it has been shown that the solubility and limited secondary structure of the two IUPs, p21 and p27, are virtually unaltered by heating to 90°C. It is anticipated that this thermal resistance to denaturation/aggregation is due to the low mean hydrophobicity and high net charge intrinsic to these proteins. This heat resistance is routinely utilized for enrichment of IDPs in mammalian cells and purification of recombinant IDPs (Galea et al. 2006)(Tsvetkov et al. 2012)(Kim et al. 2000).

1.1.3.Disorder and Diseases

Since IDPs play important biological functions and occupy central positions in protein interaction networks, many IDPs are implicated in various human diseases (V. N. Uversky, Oldfield, and Dunker 2008)(Midic et al. 2009). IDPs involved in protein interaction networks known as "hubs" associate with large numbers of partners (as mentioned earlier). In any organism, these hubs maintain normal function and stability in protein–protein interaction networks. The deletion of a hub protein can prove to be detrimental to the survival of the organism (Jeong et al. 2001). In fact, many IDPs involved in the pathogenesis of human diseases are hubs (V. N. Uversky 2011). Hubs are sensitive to the environment; they form different monomeric, oligomeric and insoluble conformations depending on the milieu.

Taking in consideration the functional roles of IDPs, Uversky and co-workers proposed that conformational diseases may originate, not only from misfolding, but also from misidentification, misregulation and missignaling (V. N. Uversky, Oldfield, and Dunker 2008). In this view, mutations and changes in environment could lead to protein confusion causing reduced ability of partner recognition in IDPs and instead lead to formation of deadly aggregates (Dunker *et al.*, 2008).

The prevalence of intrinsic disorder in proteins associated with cancer, cardiovascular disease, neurodegenerative diseases, amyloidoses, and several other human diseases were shown through the computational analysis by a number of disorder predictors on specific datasets of proteins associated with a given disease (V. N. Uversky et al. 2006)(V. N. Uversky, Oldfield, and Dunker 2008)(Iakoucheva et al. 2002)(Cheng et al. 2006)(Mohan et al. 2008)(Xue et al. 2010)(Goh, Dunker, and Uversky 2008b)(Goh, Dunker, and Uversky 2008a). Many of the analysed proteins were found to be either completely disordered or contain long disordered regions. **Figure 1.7** illustrates that intrinsic disorder is highly prevalent in CVD, diabetes, cancer, and neurodegenerative disease related proteins since their prevalence is comparable to that of signalling proteins and significantly exceeds the level of intrinsic disorder in eukaryotic proteins from SWISS-PROT and in non-homologous structured proteins from the PDB. A very significant percentage of cancer-associated (79%) and CVD-associated (61%) proteins were found to possess long regions of predicted disorder (Iakoucheva et al. 2002)(V. N. Uversky et al. 2006)(Cheng et al. 2006)(V. N. Uversky, Oldfield, and Dunker 2008). This high frequency of intrinsic disorder in disease-associated proteins gave rise to the disorder in disorders concept, popularly termed as "D^2 concept" by Uversky *et al*.(V. N. Uversky, Oldfield, and Dunker 2008).



Adapted from: BMC Genomics | 2008 | Vol.9 | (Dunker *et al*., 2008)

Figure 1.7. Abundance of intrinsic disorder in disease-associated proteins. Percentages of disease associated proteins with greater than 30-100 consecutive residues predicted to be disordered. For comparison, corresponding data for signalling and ordered proteins are shown. Analyzed sets of disease-related proteins included 1786, 487, 689, and 285 proteins for cancer, CVD, neurodegenerative disease and diabetes, respectively.

Recent examples illustrating the D^2 concept include a large set of oncogenic chimeric proteins (Hegyi, Buday, and Tompa 2009), such as the EWS-FLI1 fusion protein in Ewing's sarcoma (Erkizan, Uversky, and Toretsky 2010); tau protein whose misfolding and aggregation are linked to Alzheimer's disease and other taupathies (Narayanan et al.

2010); an abundant chromatin architectural protein, methyl-CpG binding protein 2 (MeCP2), mutations in which lead to developmental abnormalities, such as Rett syndrome (Ghosh et al. 2010)(Yang et al. 2011); myelin basic protein (MBP), dysregulations of which are associated with the autoimmune disease multiple sclerosis in humans (Homchaudhuri et al. 2010); a highly conserved tissue-specific transcriptional regulator, Limb-bud and heart (LBH) protein, aberrant gain-of-function of which is associated with partial trisomy 2p syndrome (Al-Ali et al. 2010), and many others (V. N. Uversky 2011).

1.2. Intrinsically Disordered Enzymes

The existence of intrinsic disorder can come across as an unexpected feature in enzymes. The catalytic step in an enzymatic process requires an enzyme to provide an environment in which the transition state of the chemical reaction can be reached more readily than in the absence of the enzyme (Wolfenden and J. Snider 2002)(J Benkovic and Hammes-Schiffer 2003)(Warshel et al. 2006). This demands the presence of very specific architecture of interacting molecules within their catalytic site for optimal function (Kamerlin and Warshel 2010). As a consequence, enzymes are often used as prototypical examples of the structure-function paradigm, in which the 3D structure of a protein determines its function. This requirement of a highly accurate structural arrangement of the active site was previously thought to extensively limit the occurrence of disorder in enzymes.

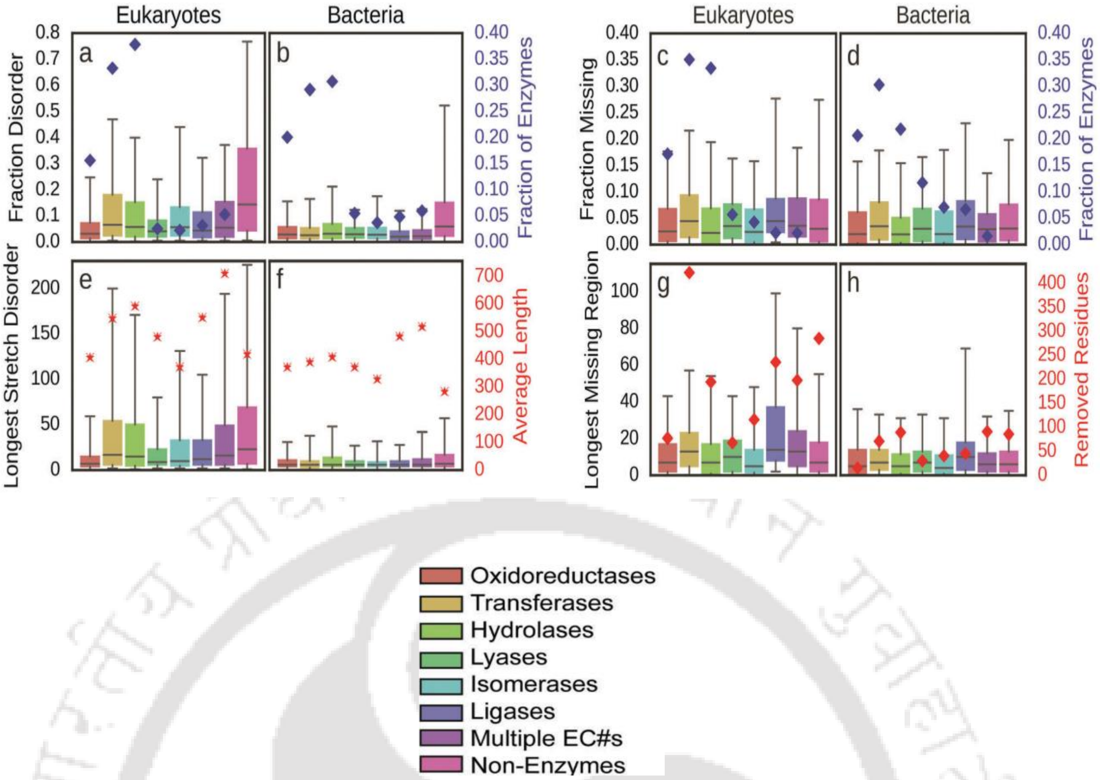
In last decade or so, a significant number of cases have been reported of enzymatic activity in disordered states. One of the first examples of a disordered protein capable of performing enzymatic activity has been that of ribonuclease T1. Its native structure contains four proline residues, two adopting the cis conformation (Tyr38-Pro39 and Ser54-Pro55) and the other two adopting the trans conformation (Trp59-Pro60 and Ser72-Pro73). The various folding intermediates of ribonuclease T1 reflect different proline isomerisation states. They possess extensive secondary structure elements and near-native topologies. Some of these intermediates have been shown to exhibit a significant fraction (40–50%) of the ribonuclease activity of the native state (Kiefhaber et al. 1992)(Aumüller and Fischer 2008).

Another work on the non-native folding intermediates of the acylphosphatase from *Sulfolobus solfataricus* reveals that they are enzymatically active. Data from stopped-flow kinetics experiments and MD simulations concluded that the catalytic site is structurally heterogeneous, whereas the remainder of the structure acts as a scaffold. In spite of being relatively disordered, the intermediate state displayed conformations in which the catalytic residues are in native-like positions. The researchers suggested that this structural organisation in the presence of the substrate creates an environment in the active site suitable for efficient catalysis (Bemporad et al. 2008).

Other noteworthy examples include a circularly permutated dihydrofolate reductase from *Escherichia coli* (V. Uversky et al. 1996) and a double mutant of a staphylococcal nuclease (Li and Jing 2000). The molten globule states of both these two proteins were found to be enzymatically active.

The molten globules are generally viewed as non-functional protein folding intermediates with near native secondary structure yet highly dynamic tertiary form with mobile side chains and a loosely packed protein core (Ohgushi and Wada 1983).The molten globule was initially discovered as an equilibrium structure observed in studies on protein denaturation. As mentioned in the above examples, there exists now enough proof in the scientific community of the fact that molten globules can carry out biological functions.

In 2017, Shelly DeForte and Vladimir N. Uversky pioneered to quantifying the extent of functionally significant intrinsically disordered protein regions (IDPRs) in enzymes. They conducted a multilevel computational analysis of missing regions in X-ray crystal structures in the PDB and predicted disorder in 66 representative proteomes. As expected, they found that the fraction of predicted disorder was higher in non-enzymes than enzymes, since non-enzymes were more likely to be fully disordered. But classes of enzyme viz. transferases, hydrolases and enzymes with multiple assigned functional classifications displayed similarity with non-enzymes in terms of the length of the longest continuous stretch of predicted disorder as shown in **Figure 1.8**. In eukaryotes, both enzymes and non-enzymes were shown to exhibit greater disorder content than that observed in bacteria (DeForte and Uversky 2017).



Adapted from: Mol. BioSystems | 2017 | Vol.13, 463-469 | (DeForte and Uversky 2017)

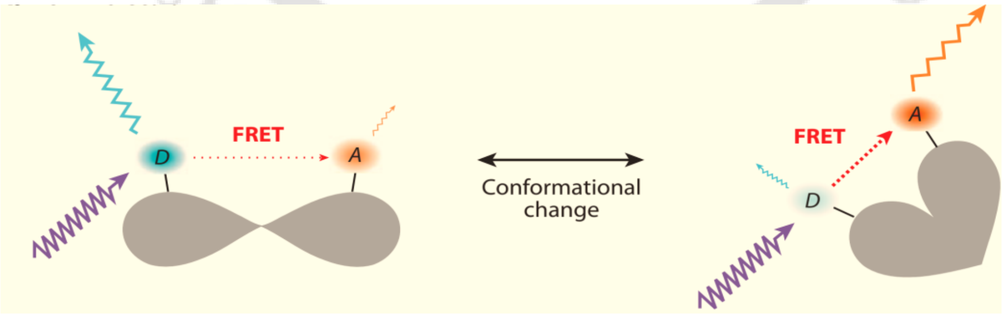
Figure 1.8. Measurements of predicted disorder in reference proteomes and measurements of missing regions in the PDB. Box plots are first quartile, second quartile (median), and third quartile. Plot whiskers extend to 5% and 95% of the data. (a and b) The fraction of predicted disorder per-protein in reference proteomes in eukaryotes (a) and bacteria (b). (c and d) The fraction of missing residues per-protein in the PDB in eukaryotes (c) and bacteria (d). (a–d) The blue diamond shows the relative fraction of the full set of enzymes that each enzyme type occupies. (e and f) The longest predicted missing region per-protein in reference proteomes in eukaryotes (e) and bacteria (f). The red star shows the average protein length. (g and h) The longest missing region per-protein in eukaryotes (g) and bacteria (h). The red diamond shows the average number of removed residues per-protein before crystallization.

Recent evidences also suggest that disordered proteins could display catalytic activity by becoming structured in the region housing the active sites, yet retain a significant degree of conformational heterogeneity. Nevertheless, there exist two main challenges for conducting optimal catalysis in a disordered state. Firstly, the requirement to establish a catalytic environment by quick conformational search of the disordered state in a time frame less than that of the overall turnover time of the enzyme. Secondly, to be able to maintain this catalytic environment long enough for catalysis to occur, in spite of its inherit structural fluctuations. These conditions can be difficult to satisfy for most proteins and hence existence of disordered enzymes is not very commonplace

(Vendruscolo 2010). The study of these rare disordered proteins can increase our general understanding of the mechanism of enzymatic catalysis and enable us to manipulate the unique attributes of these proteins to our advantage.

One crucial factor to consider in order to understand how these intrinsically disordered enzymes operate is the functional dynamics of the protein. A comparative and quantitative study of structure, dynamics and activity of two homologs (hyper-thermophilic and mesophilic) of adenylate kinase revealed a close link between the enzyme dynamics and catalytic turnover. They measured the residue-specific protein dynamics and catalytic turnover by NMR relaxation studies and found that the nucleotide binding lids in this enzyme, a dynamic process, was rate-limiting for both homologs (Wolf-Watz et al. 2004). Therefore, the catalysis by an enzyme is not only defined by the actual chemical steps but also by its structure and dynamics.

To monitor the conformational dynamics of protein in native and functional states, Fluorescence Resonance Energy Transfer (FRET) has emerged as a powerful, reproducible, sensitive and site-specific tool (Stryer 1967)(Ma, Yang, and Zheng 2014). For example, the local conformational changes after binding a suitable ligand/substrate in the vicinity of a FRET pair in a protein can be interpreted in terms of the deviations in distance between this FRET pair as a result of increase/decrease of FRET efficiency between the donor and acceptor pair as depicted in **Figure 1.9**. (refer Chapter 2, FRET section for details)

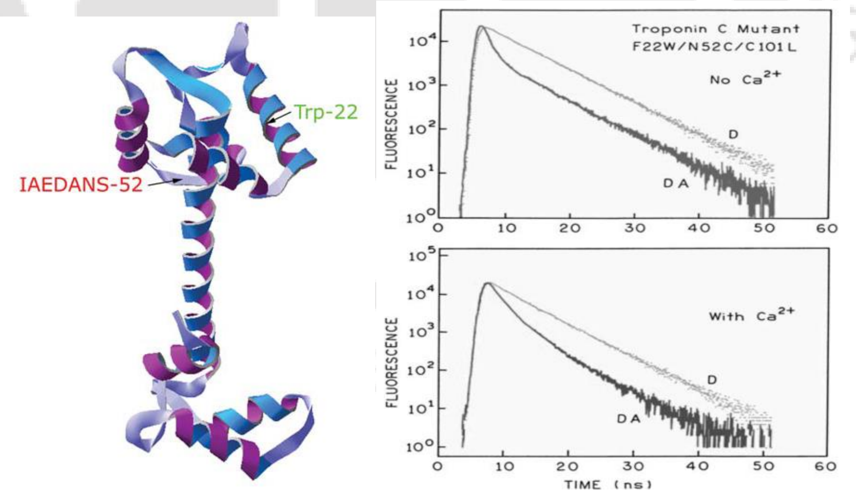


Source: Annual Reviews Biochemistry | 2011 | Vol. 80 | (Miyawaki 2011)

Figure 1.9. FRET (fluorescence resonance energy transfer) within a protein molecule. D, donor; A, acceptor. Zigzag arrows of violet, light blue, and orange indicate excitation light for the donor, emission from the donor, and emission from the acceptor, respectively.

The intrinsic fluorophore in proteins, tryptophan (trp), is a very convenient probe to detect local transitions in protein structure and dynamics (Ghisaidoobe and Chung 2014). The high sensitivity of tryptophan to its local environment (solvent/amino acid residues) further makes it a reliable tool to assess the subtle changes in the immediate environment surrounding the tryptophan moiety (Vivian and Callis 2001). A native protein containing tryptophan as a fluorescence donor and an extrinsic probe labeled to another suitable amino acid (usually cysteine/lysine) as the fluorescence acceptor can act as a system to monitor any transition in structure and dynamics accompanying the protein's switch to a function state through FRET analysis (J. R. Lakowicz et al. 1994)(J. Lakowicz 2006)(Gustiananda et al. 2004).

For example, the Ca^{2+} sensitivity of a muscle contraction protein Troponin C, a tryptophan less protein, was studied using a combination of FRET and site-directed mutagenesis (**Figure 1.10**). A FRET pair was constructed by inserting a single trp residue at position 22 to serve as the donor and a cysteine (cys) residue at position 52 to act as a reactive site for the acceptor fluorophore IAEDANS [N-(Iodoacetaminoethyl)-1-naphthylamine-5-sulfonic acid]. Time-dependent FRET was used to determine the variations in donor-acceptor distance in Ca^{2+} unbound and bound states (She et al. 1998)

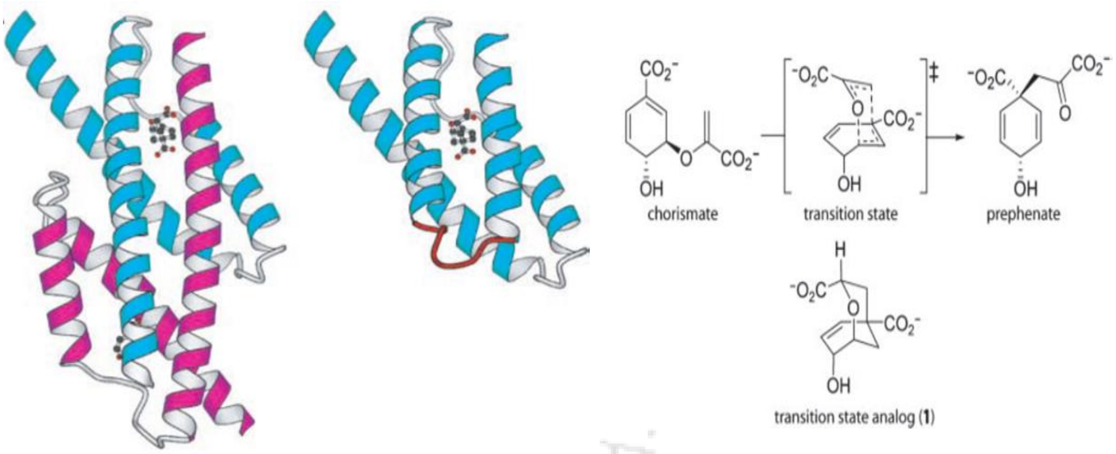


Source : J Mol Biology | 281:445–452 | (She et al. 1998) and Principles of Fluorescence Spectroscopy| 3rd Edition | (J. Lakowicz 2006)

Figure 1.10. Turkey skeletal 2Ca–Troponin C. *(Left)* The crystal structure *(Right)* Tryptophan-22 fluorescence lifetime intensity decays of Troponin C without acceptor (D) and with an acceptor (DA), an IAEDANS on cysteine 52. *Top panel*: without Ca^{2+} and *Bottom panel*: with Ca^{2+}.

For this thesis work, we decided to adopt a similar strategy based on site-directed mutagenesis (SDM) coupled with FRET based methods to study the structure-function relationship of an intrinsically disordered enzyme, monomeric chorismate mutase (mMjCM). mMjCM or simply referred to as mCM (in this thesis) contains two tryptophan residues at positions 24 and 26 as well as a single cysteine residue at position 69. This feature makes it a very suitable candidate to monitor the changes in enzyme structure and dynamics in response to a substrate/ligand utilizing FRET between the Trp (24/26) and Cys (labeled with a fluorescence acceptor probe, dansyl) (Bhatia, Krishnamoorthy, and Udgaonkar 2018)(Hammarstro and Persson 2001). Moreover, by introducing this single Cys residue at different sites on the protein via amino acid replacement (Cys scanning site-directed mutagenesis), a better understanding of how different segments of mCM come together during catalysis can be established. Details of the sites selected for mutagenesis and the logic behind the selection process will be discussed in Chapter 3.

Monomeric Chorismate Mutase (mCM): mCM is a 109 amino acid residues protein obtained by the topological redesign of the highly active enzyme, AroQ chorismate mutase, from *Methanococcus jannaschii* (MjCM). A process of genetic selection in combination with structure-based design was utilized to convert the dimeric helical bundle enzyme (MjCM) into a monomeric four-helix-bundle protein (mMjCM), while still retaining its enzymatic activity (MacBeath, Kast, and Hilvert 1998). This structural re-design was achieved by inserting the hinge-loop sequence ARWPWAEK (loop positions referred to as 23a-h by the Hilvert group) into the long, dimer spanning N-terminal helix of the dimeric naturally occurring MjCM (MacBeath, Kast, and Hilvert 1998) as shown in **Figure 1.11**.



Source: PNAS | August 31, 2004 | vol. 101 | no. 35| (K. Vamvaca et al. 2004)

Figure 1.11. Topological redesign of the enzyme MjCM. (*Left*) The thermostable MjCM homodimer (Left) was converted into a monomer (mMjCM, middle) by inserting a flexible hinge loop (red) into the long H1 helix. The models are based on the X-ray structure of a related *E. coli* chorismate mutase domain complexed with a transition-state analog (TSA), which is shown in in ball-and-stick representation in the model. (*Right*) Both enzymes efficiently catalyze the rearrangement of chorismate to prephenate.

The active mCM protein catalyzed the conversion of chorismate to prephenate, as done by its natural dimeric parent enzyme (MjCM). This catalytic conversion is a crucial step in the biosynthesis of aromatic amino acids. The parameters k_{cat} and K_m for both these enzymes were found to be comparable (k_{cat} 3.2 s^{-1}, K_m 170 µM), thus implying their near identical functional efficiency (Vamvaca *et al.*, 2004)(MacBeath, Kast, and Hilvert 1998).

The active site of monomeric chorismate mutase is depicted in **Figure 1.12**. The network of electrostatic and hydrogen bonding networks between the transition state analog and the active site residues is indicated in the figure. The numbering of residues is based on EcCM as the active site closely resembles the *E.coli* chorismate mutase (Lee et al. 1995, Vamcava et al.2008). Overall, there are 12 hydrogen bonds or electrostatic interactions between the enzyme and the ligand (TSA) and every oxygen atom of the TSA is involved in two polar interactions. The TSA has no solvent-accessible surface area in the complex and is completely buried. A tightly bound water bridging the two carboxylates and Arg 51 is present.

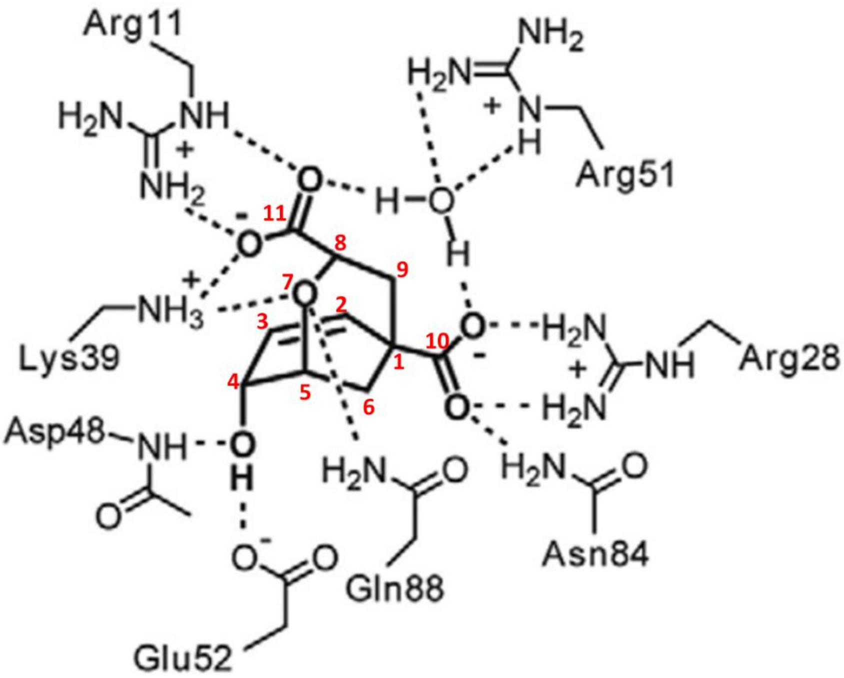


Figure 1.12. Active site of monomeric chorismate mutase (sequence numbering based on EcCM) [Adapted from Vamvaca et al.2008]. The ligand (TSA) atoms are numbered 1 to 11 in red. The hydrophobic active-site residues that are involved in van der Waals contacts with the apolar portions of the ligand are not represented here.

The fact that mCM could retain the functional attributes of its parent enzyme came as a surprise to the concerned investigators. This is due to the fact that mCM possessed biophysical properties associated with a non-native molten globule state, viz. poor NMR signal dispersion, rapid hydrogen-deuterium exchange, binding to 8-anilino-1-naphthalene sulfonate (ANS) and noncooperative thermal denaturation (Ptitsyn et al. 1990)(Dolgikh et al. 1981)(Kuwajima 1989)(Betz, Raleigh, and DeGrado 1993)(Ohgushi and Wada 1983).

In NMR analysis, the [^{15}N, ^{1}H]-TROSY spectrum of the monomer mCM displayed properties of a poorly packed protein with multiple energetically accessible conformations. The observed peaks were poorly dispersed and the number of signals obtained were much less than that expected for a native-like protein. In a disordered state, the rapid interconversion of many energetically accessible conformations lead to the poor dispersion of resonances (Vamvaca *et al.*, 2004). Interestingly, after addition of a ligand, a transition state analogue (TSA), mCM exhibited a well resolved NMR spectrum typical of a folded protein. This was a clear indication of a disorder to order transition of mCM in presence of a suitable ligand (Vamvaca *et al.*, 2004).

A substantial increase in fluorescence intensity as well as a characteristic blue shift of the emission spectra was observed for the hydrophobic dye ANS bound to mCM, as expected from a molten globule (Demarest et al. 1999) (Semisotnov et al. 1991). ANS can also bind to mCM to some extent in presence of the ligand TSA. This suggests that mCM, in ligand bound state, can still retain considerable dynamic character (Vamvaca *et al.*, 2004).

H/D exchange experiments added to the proof of high conformational flexibility in mCM (Englander et al. 2008). Electrospray ionization MS was used to monitor the rate of exchange. Merely after initiation of the exchange (10s), most of the labile protons in mCM were deuterated. A significant fraction (35%) of the protons could be protected from exchange initially by the addition of TSA, but within 5 mins, the mCM-TSA complex reported a degree of deuteration comparable to that of the unbound enzyme. These results indicated that mCM retained a highly mobile structure even in the ligand bound state (Vamvaca *et al.*, 2004).

It was suggested by Hilvert et al. that the intrinsically disordered molten globular nature of ligand unbound mCM arose as a consequence of the absence of a well packed hydrophobic core. The active site in the interior of mCM was found to be highly polar (Lee et al. 1995) and this high density of charged residues at the core of the protein might have prevented the formation of a stable native-like folded state in absence of ligand. However, a ligand/substrate can fill up this core, providing a template around which the protein can start "crystallizing". Furthermore, improved packing throughout the protein structure can then be achieved by the extensive H-bonding and electrostatic interactions between ligand/substrate and mCM (Vamvaca *et al.*, 2004).

Interestingly, the energy cost for the enhanced flexibility of mMjCM in comparison to MjCM is very small. The catalytic efficiency (k_{cat}/K_m) of mCM and the dimer MjCM differ only by a factor of three (approximately 0.6 kcal/mol), due to their differences in K_m (MacBeath, Kast, and Hilvert 1998). Some preliminary calorimetric measurements show that the unfavourable entropy associated with preorganizing the enzyme is counteracted by enthalpic gains from the noncovalent interactions between the TSA and mCM (Vamvaca *et al.*, 2004). This finding was in accordance with the proposal by

William et al. that catalytic efficiency can originate from improved packing interactions in the transition state (D. H. Williams, Stephens, and Zhou 2003). The structural ordering seen in case of mCM can be considered as an extension of the classic induced-fit model of enzyme catalysis (Daniel E. Koshland 1995) in which the active site structural rearrangement by the substrate lead to a high affinity complex. mCM represents an extreme case of this classical model, where structural rearrangement is not just confined to local structuring but in fact induces a global ordering (Vamvaca *et al.*, 2004).

DeGrado in his Nature article discussed that many if not most modern-day enzymes could have evolved from molten globules (DeGrado 1993). In a primitive cell in the primordial world, the structural plasticity of a molten globule enzyme which confers relaxed substrate specificity and/or catalytic promiscuity (James and Tawfik 2003) would have provided significant evolutionary advantage, allowing the performance of several functions in the primitive cell (Katherina Vamvaca, Jelesarov, and Hilvert 2008) (Vamvaca *et al.*, 2004). mCM holds a strategic position in designing a variety of novel structures and functions by taking advantage of its inherent conformational diversity and channelling it down multiple evolutionary trajectories by applying random mutagenesis and powerful selection or screening technologies (Vamvaca *et al.*, 2004).

1.3. Key questions addressed in the thesis

- What is the nature of structural plasticity mechanisms in engineered monomeric chorismate mutase (mCM) that enable catalysis?

- How do designed mutations alter catalysis? How are the segmental dynamics, conformational adaptation, and disorder-to-order transition correlated with changes in catalysis?

With the following specific aims

- Generation of single tryptophan-cysteine pair mutants in mCM through site-directed mutagenesis
- Expression and purification of wild-type mCM (wt mCM)and its mutants
- Comparison of the biochemical (enzyme activity) and biophysical properties of mCM and its mutants
- Monitoring changes in dynamics of the wt mCM and its mutants in the presence of suitable ligands using fluorescence spectroscopy

CHAPTER 2

Experimental Techniques, Materials and Methods

2.1. Experimental techniques

The principles of some of the important molecular biology and biophysical techniques used throughout this thesis have been explained here in brief.

2.1.1. Site-directed mutagenesis

In vitro site-directed mutagenesis is a powerful technique for characterizing the dynamic, complex relationships between protein structure and function. Several approaches to this technique have been published, but these methods generally utilize single-stranded DNA (ssDNA) as the template (Nelson and McClelland, 1992)(Papworth, C. Bauer and Braman, 1996). This can be quite labour intensive or technically complicated. Therefore, we resorted to the use of QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies) which allows site-specific mutation in virtually any double-stranded plasmid. Hence the conventional process of DNA subcloning and ssDNA rescue is no longer required. Moreover, this simple mutagenesis technique does not require specialized vectors, unique restriction sites, multiple transformations or in vitro methylation treatment steps. A rapid three-step procedure generates mutants with greater than 80% efficiency in a single reaction. The simple protocol uses either miniprep plasmid DNA or cesium-chloride-purified DNA.

The basic procedure utilizes a supercoiled double-stranded DNA (dsDNA) vector with an insert of interest and two synthetic oligonucleotide primers, both containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, are extended during temperature cycling by PfuUltra HF DNA polymerase, without primer displacement. Extension of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. Following temperature cycling, the product is treated with Dpn I. The Dpn I endonuclease (target sequence: 5´-Gm^{6}ATC-3´) is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template [DNA isolated from almost all *E. coli* strains is methylated by deoxyadenosine methylase (dam) and therefore susceptible to Dpn I digestion]. Consequently, only mutation-containing synthesized DNA remains in its intact form (Nelson and McClelland, 1992). The nicked vector DNA containing the desired mutations is then transformed into XL1-Blue super-competent cells or DH5α competent cells. The confirmation of the incorporation of desired mutations in the plasmid vector of interest can then be done by isolation of plasmid DNA from positive bacterial colonies followed by DNA sequencing.

2.1.2. Molecular spectroscopy

Spectroscopy is a branch of science which identifies and measures interaction of electromagnetic radiation (ER) with matter. These interactions can be ascribed by absorption, emission or scattering. The light induced transition of matter from a lower energy state to a higher energy state is the basis of absorption spectroscopy. The molecular transition strongly depends upon the frequency of the radiation absorbed. For instance, X-ray radiation which oscillates at a high frequency is energetic enough to excite core electrons, whereas lower frequency ultraviolet and visible radiation only excites valence electrons. Lowest energy transitions occur in the vibrational and rotational levels of molecules and nuclei and are induced by infrared and radio waves. Emission spectroscopy, on the other hand, deals with the relaxation of an atom or molecule from a higher energy state to a lower one. This transition is usually accompanied by the emission of radiant energy which can be measured by a variety of spectroscopic techniques specific to the frequency of the radiation. The various processes associated with the absorption and emission of radiation can be illustrated by Jablonski diagram (**Figure 2.1**).

Subsequent to the absorption of light, electrons are excited from the lowest energy level S_0 to higher energy levels either S_1 or S_2. The energy difference between the two states is equal to the energy of the incident photon. The electronic transition occurs very fast with the time scale of 10^{-15} seconds which is much faster than nuclear reorganization. Therefore, the higher energy states are achieved without any apparent change in the position of nuclei. This hypothesis is known as Franck-Condon principle. After absorption, a molecule returns to the ground state by radiative processes like fluorescence or phosphorescence and non-radiative processes such as vibrational relaxation, internal conversion, intersystem crossing, fluorescence quenching and other deactivation processes.

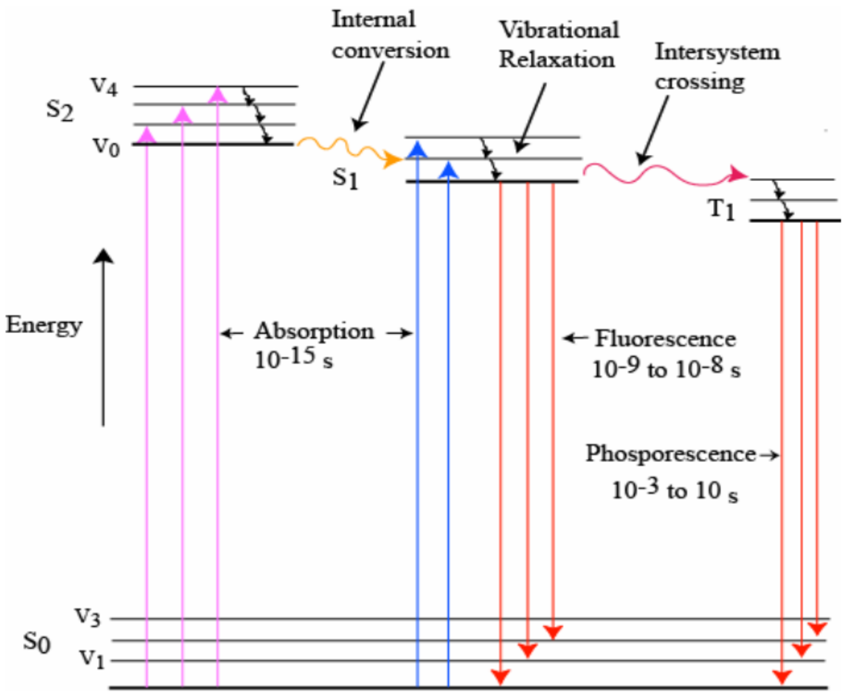


Figure 2.1. Jablonski diagram [Source: (Lakowicz, 2006)]

2.1.2.1. Absorbance

Absorptivity is the basic property of a chromophore and is constant at a particular wavelength. Absorptivity is the basic property of a chromophore and is constant at a particular wavelength. When a photon of light encounters a chromophore possessing quantized electronic states, it may be absorbed and its energy transferred to an electron to partake in an electronic transition described by:

$$E_{\text{photon}} = \Delta E_{\text{electronic transition}} = h\nu$$

In classical electronics, the energy due to interaction of an electric field E and a dipole is given by the dot product of the dipole moment and the electric field vector.

$$\text{Energy} = \vec{E} \cdot \vec{\mu} \qquad 2.1$$

The dipole moment is defined as the summation of the product of the charge q_j times the position vector r_j for all charged particles j.

$$\mu = \sum_j q_j r_j \qquad 2.1.1$$

State of molecules changes subsequent to their interaction with light. For example, molecule with initial state ‘a’ changes to state ‘b’ subsequent to their interaction with light. This perturbation in states is outcome of altered distribution of electric charge of molecule in the presence of oscillating electric field E. The transition probability from state ‘a’ to ‘b’ is governed by transition dipole moment $\langle\psi_b|\tilde{\mu}|\psi_a\rangle$. abbreviated as μ_{ba}. Here Ψ_b and Ψ_a are the wavefunction of two states b and a. The absorption intensity depends upon the rate at which energy is taken up from incident light. Transition rate (dP_b/dt) is product of two terms and can be written as

$$dP_b/dt = B_{ab}I(\nu) \qquad 2.2$$

Here, B_{ab} is the transition rate per unit energy density and $I(\nu)$ denote the energy density incident on the molecule. B_{ab} can be written as

$$B_{ab} = (2/3)(\pi/\hbar^2)|\langle\psi_b|\tilde{\mu}|\psi_a\rangle|^2 \qquad 2.3$$

The molar extinction coefficient is related to Einstein coefficient for stimulated absorption as

$$B_{ab} = (1000c/N_0\,h)\int(\varepsilon'/\nu)\,d\nu \qquad 2.4$$

where N_0 is Avogadro’s number, c is the velocity of the light and ν is frequency of light.

Using the equation 2.3 and 2.4, and converting ε' to ε, the following relationship can be obtained

$$|\langle\psi_b|\tilde{\mu}|\psi_a\rangle|^2 = 9.180\ \times 10^{-3}\int(\varepsilon'/\nu)\,d\nu = D_{ab}\ (\text{debye}^2) \qquad 2.5$$

Here D_{ab} is the dipole strength, which is a measure of the strength of transition dipole. Apart from this, another useful term is oscillator strength, f_{ab} which compares the intensity of absorption to that expected from a three-dimensional harmonic oscillator, given by

$$f_{ab} = (8\pi^2 mc/3h\nu)D_{ab} = 4.315\times10^{-9}\int \varepsilon(\nu)d\nu \quad (\text{dimensionless}) \qquad 2.6$$

where m is the mass of electron. For strongly allowed transitions f_{ab} are in the range of 0.1 to 1. Both the quantities, f_{ab} and D_{ab} are crucial in understanding the various special optical effects observed in polymers.

When a radiation from given light source falls on solution containing a chromophore, it absorbs the light to a certain extent. The relationship between the amount of light absorbed and the concentration of the chromophore is expressed by Beer-Lambert law:

$$A = log_{10}(I_0/I) \qquad 2.7$$

where A is the absorbance, I_0 is the intensity of the incident light, I is the intensity of the transmitted light. The absorbance is linearly related to the concentration of the solution which can be expressed as

$$A = \varepsilon(\lambda).c.l \qquad 2.8$$

where ε is the molar extinction coefficient at given wavelength, c is the molar concentration of the absorbing molecule and l is the path length travelled by the light in solution.

2.1.2.2. Fluorescence

Fluorescence is a process in which emission of visible or ultraviolet light accompanying a transition of electrons from excited singlet state (S_1) to ground singlet state (S_0) occurs. In the fluorescence, emitted light has longer wavelength compared to absorbed light because of energy loss via vibrational relaxation and other non-radiative processes. Properties of biological molecules like its dynamics and interaction with other molecules revealed by emitted light is quite different in comparison with properties revealed by absorption of light. The reason for this is, the emission process is occurring at much slower timescale (~ nanosecond) than absorption (~femtosecond). Due to slower timescale, emission spectrum is influenced by perturbations and wide range of molecular interactions.

Factors that affect fluorescence intensity:

A. Internal Conversion:

In this process, excitation energy is lost by collision with solvent or by dissipation through internal vibrational modes. These types of transitions take place on the order of 10^{-12} seconds, and is represented in **Figure 2.1** by an orange wavy arrow between S_2 with S_1. In general, rate of internal conversion (k_{ic}) has been observed to increase with increasing temperature. Therefore, the observed fluorescence decreases with the increase in temperature. Monitoring macromolecular conformation changes induced by heat becomes complicated because of this intrinsic temperature dependence.

B. Intersystem crossing:

Molecules in the S_1 state can also undergo a spin conversion to the first triplet state T_1. Conversion of S_1 to T_1 is called as intersystem crossing and is depicted as magenta wavy arrow (k_{isc}) in **Figure 2.1**. Excited triplet state T_1, in turn, convert to ground state either by phosphorescence or by internal conversion. Since transition from T_1 to singlet ground state is forbidden, rate constant for triplet emission are several orders of magnitude smaller than those for fluorescence.

Due to occurrence of non-radiative process (internal conversion and intersystem crossing) at the same time scale as fluorescence, the observed fluorescence lifetime (τ_F) can be defined as

$$\tau_F = \frac{1}{k_r + k_{isc} + k_{ic} + k_q[Q]} \qquad 2.9$$

Where k_r, k_{isc}, k_{ic} and k_q denote the rate of decay by fluorescence, intersystem crossing, internal conversion and fluorescence quenching respectively. [Q] is the quencher molar concentration.

The above equation can be simplified as

$$\tau_F = \frac{1}{k_r + k_{nr}} \qquad 2.10$$

k_r and k_{nr} represents rate constant of radiative decay and rate constant of all non-radiative decay, respectively.

The quantum yield (ϕ_F) or fluorescence efficiency can be expressed as follow

$$\phi_F = \frac{k_r}{k_r + k_{isc} + k_{ic} + k_q[Q]} = \frac{k_r}{k_r + k_{nr}} \qquad 2.11$$

Fluorescence lifetime in absence of non-radiative process is known as intrinsic or natural lifetime (τ_n) and is a reciprocal of k_r. Theoretically, τ_n can be calculated from absorption spectrum.

$$\tau_n = \frac{1}{k_r} \qquad 2.12$$

Relationship between fluorescence lifetime and quantum yield can be defined as

$$\phi_F = \frac{\tau_F}{\tau_n} \qquad 2.13$$

Interdependence between quantum yield and fluorescence lifetime suggest that decrease in the fluorescence lifetime results in decreasing quantum yield.

2.1.2.2.1 Steady-state fluorescence

Measurements of fluorescence can be broadly categorized into two types: steady-state and time-resolved measurements. In steady-state fluorescence measurements, sample is irradiated with continuous beam of light and subsequently fluorescence intensity or emission is recorded in time scale of seconds. In steady state, equilibrium is established between the population of fluorophore in the excited state and ground state. This method is used to measure spectral distribution of fluorescence intensity and also can be used to calculate quantum yield of fluorophore. Steady state fluorescence method can be applied to study different fluorescence phenomena like steady state anisotropy, resonance energy transfer, quenching etc. However, the steady state fluorescence intensity gives a time averaged information of fluorescence dynamics. Variations in fluorescence intensity and emission maximum can be utilized to studies many important biological phenomena. Fluorescence spectrum can also reveal information about the solvent exposure of fluorophore and degree of polarity around fluorophore.

2.1.2.2.2 Steady-state fluorescence anisotropy

Fluorescence anisotropy technique gives information about the rotational motion of fluorophore in the excited state which depends on shape and size of fluorophore. Anisotropy is utilized to study protein-protein interaction, protein aggregation, protein folding and membrane fluidity. In fluorescence anisotropy, sample is illuminated with vertically polarized light and sample may emit polarized or unpolarized light. Fluorescence anisotropy (r) is measured by monitoring the extent of polarization of emission.

Steady-state anisotropy (r_{ss}), is an average of anisotropy decay time weighted over the intensity decay of the fluorophore and reveals the information about the overall rotational motion of molecules.

$$r_{ss} = \frac{\int_0^\infty r(t)I(t)dt}{\int_0^\infty I(t)dt} \qquad 2.14$$

Intrinsic or extrinsic probe attached with macromolecules shows both local and global rotational motion due to tumbling of the macromolecule through Brownian motion. Depolarization of excited state occurs because the tumbling of molecule in solution.

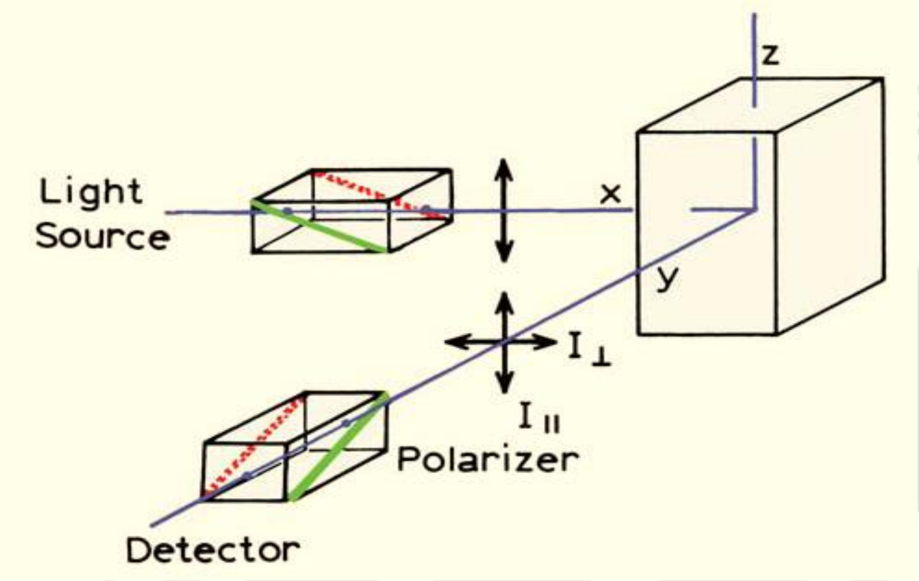


Figure 2.2 Image above showing the schematic of fluorescence anisotropy (Source: Principle in Fluorescence Spectroscopy by J. R. Lakowicz, third edition, 2006).

Fluorophores in the homogeneous solution are randomly oriented in their ground state at room temperature. When samples are excited with polarized light, molecules have the higher probability of excitation if its absorption transition dipole are oriented parallel to the electric vector of the excited light. This process is termed as photoselection. In the period of fluorophore excited state lifetime (~ 10^{-9} s), extent of polarization changes because of Brownian rotational motion of molecule. This extent of depolarization depends on rate of rotational diffusion and can be measured in term of anisotropy. If fluorophore has fluorescence lifetime comparable to its rotational rate it shows non-zero anisotropy. Anisotropy value is close to zero for fast tumbling fluorophore.

Anisotropy can be expressed by the following equation

$$r_{ss} = \frac{I_{\parallel} - GI_{\perp}}{I_{\parallel} + 2GI_{\perp}} \qquad 2.15$$

Herer_{ss} denotes steady-state anisotropy, $I_{\parallel}$ represents the fluorescence emission intensity when emission and excitation polarizers are vertically aligned and $I_{\perp}$ denotes the fluorescence emission intensity when emission polarizer aligns perpendicular to the excitation polarizer. Schematic of fluorescence anisotropy measurement is shown in **Figure 2.2**. G is known as geometry factor (G-factor) and defined as the ratio of monochromator and detector sensitivity to intensity of vertically (S_V) and horizontally (S_H) polarized light.

$$G = \frac{S_V}{S_H} \qquad 2.16$$

The value of G factor depends on the emission wavelength and bandpass of the monochromator. Anisotropy is a dimensionless quantity and does not depend on the concentration of fluorophore in absence of any artifacts, like scattering.

2.1.2.2.3. Time-resolved fluorescence decay

In time-resolved fluorescence using time-correlated single photon counting (TCSPC) method, sample is excited with a short duration (nanoseconds or less) pulse of light and subsequently decay of fluorescence intensity is analyzed as a function of time. Time-resolved measurements can be studies by two methods such as time domain method and frequency domain method. In time domain, sample is excited with pulse of light for ultra-short duration. Whereas in frequency domain sample is excited by intensity modulated light. In this thesis, we have utilized the time domain method (TCSPC) for our lifetime fluorescence measurements.

In this method, sample is illuminated by a pulse of light with a very short temporal pulse width typically in sub-nanosecond or picosecond range. The pulse width should be as short as possible and ideally far shorter than decay time (τ) of fluorophore.

After exciting the sample with short pulse of light, time-dependent fluorescence intensity decay of fluorophore is measured using a technique known as Time Correlated Single Photon Counting (TCSPC). From the slope of log$I(t)$ versus t plot, fluorescence decay time (τ) is measured. To avoid the effects of anisotropy or rotational diffusion on the fluorescence intensity decay, the fluorescence intensity decay is generally collected at 54.7° through a polarizer. Set up for fluorescence time-resolved measurements is shown **Figure 2.3**.

If the fluorophore has single lifetime, intensity decay equation can be expressed as

$$I(t) = I_0 exp\,(-t/\tau) \qquad 2.17$$

Here $I(t)$ denotes intensity at any time t, I_0 denotes the initial intensity of fluorophore and τ is the fluorescence lifetime of fluorophore.

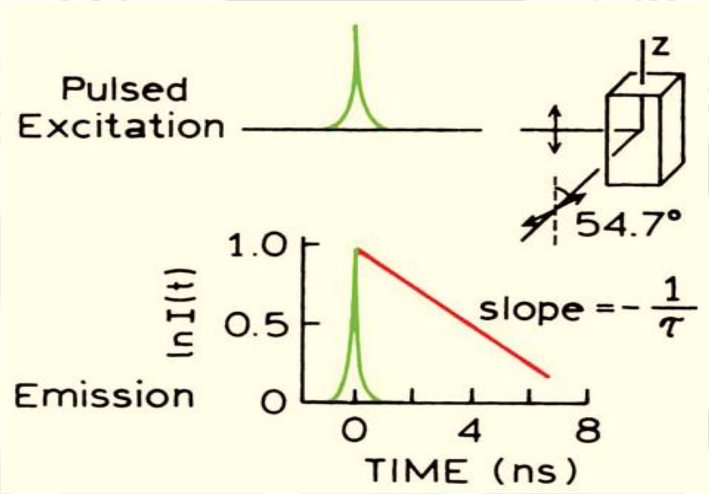


Figure 2.3. Schematic for Time-domain fluorescence lifetime measurements (Adapted from Principle in Fluorescence Spectroscopy by J. R. Lakowicz, third edition, 2006).

The fluorescence lifetime is a measure of average amount of time a fluorophore stays in its excited state before emission of a photon. Fluorescence lifetime for a single exponential decay can be measured by plotting the fluorescence intensity decay of the sample as a function of time (t) as shown in **Figure 2.3**.

Fluorescence lifetime (τ) can be defined as follows:

$$< t > = \frac{\int_0^\infty tI(t)dt}{\int_0^\infty I(t)dt} = \frac{\int_0^\infty t \exp\,(-t/\tau)dt}{\int_0^\infty \exp\,(-t/\tau)dt} \qquad 2.18$$

By solving above equation, the average time of a fluorophore stay in the excited state is measured to be equal to fluorescence lifetime τ.

$$< t > = \tau \qquad 2.19$$

The sample shows multi-exponential decay if fluorophore present in different environments or due to presence of more than one fluorophore. Time-dependent intensity for a multi-exponential decay can be expressed as

$$I(t) = \sum_{i} \alpha_i \, exp \, (-t/\tau_i) \qquad 2.20$$

Here τ_i and α_i represents the ith fractional lifetime and ith fractional amplitude respectively for the decay component. Addition of fractional amplitude is equals to unity and can be written as

$$\sum_{i} \alpha_i = 1 \qquad 2.21$$

By using the above equation, mean lifetime τ_m can be measured as

$$\tau_m = \sum_{i=1}^{n} \alpha_i \, \tau_i \qquad 2.22$$

τ_m is directly proportional to area under the intensity decay curve. As steady-state measurement gives an average of the time-resolved event over the fluorescence intensity decay of the fluorophore, it is important to know the relationship between steady-state and time-resolved fluorescence.

If a fluorophore displays a single decay time *(τ)* its intensity decay can be described as

$$I(t) = I_0 \, e^{-t/\tau} \qquad 2.23$$

Relationship between decay time and steady-state intensity (I_{ss}) can be describe as

$$I_{ss} = \int_{0}^{\infty} I_0 \, e^{-t/\tau} dt = I_0 \, \tau \qquad 2.24$$

Time-resolved measurement has many advantages in comparison with steady-state. Some important advantages of time-resolve method are: 1) it does not depend on the fluorophore concentration, 2) tells about the microenvironment around fluorophore by revealing multiple lifetimes and 3) explain about the type of quenching like Static or dynamic quenching.

2.1.2.2.4 Time-resolved fluorescence anisotropy decay

Time-resolved fluorescence anisotropy technique is used to reveal the rotational dynamics of the molecule. Time-resolved anisotropy gives information about two types of rotational motion in a molecule: 1) segmental or localized motion, arising due to the independent rotation of fluorophore attached covalently to macromolecule with fast dynamics and a short rotational correlation time and 2) global motion, arising due to the tumbling of whole macromolecule in the solution with a slow dynamics and a long correlation time (**Figure 2.4**)

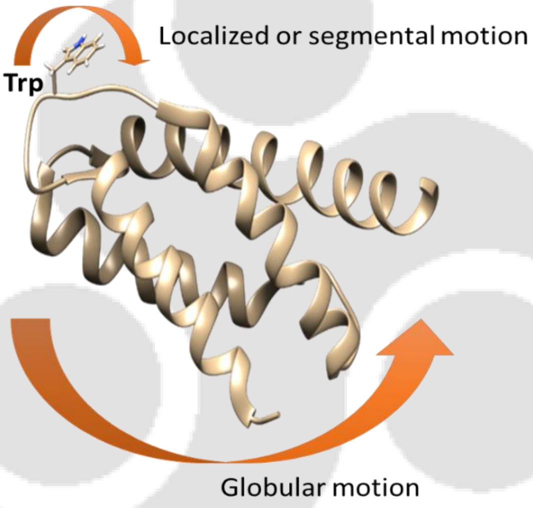


Figure 2.4. Illustration of the global and localized (Trp probe) motion in the protein mCM.

Fast rotational motions in macromolecules can be studied by time-resolved fluorescence and depolarization on the time scales of nanosecond and sub-nanosecond. Information about shape, size and flexibility of macromolecules can be revealed by time-dependent anisotropy *r(t)*

After exciting the sample with short pulse of light, the anisotropy decay of spherical body rotating in the solution is given by

$$r(t) = r_0\, e^{-t/\theta} = r_0\, e^{-6Dt} \qquad 2.25$$

Here r_0 is initial anisotropy at time (t) = 0, D denotes as the rotational diffusion coefficient and θ indicates the rotational correlation time of molecules rotating freely in the solution. The value of rotational correlation time (θ), depends on temperature (T), viscosity (η) and hydrodynamic volume (V) of the molecule. Rotational correlation time of a sphere (θ), can be explained by the Stokes-Einstein equation.

$$\theta = \frac{\eta V}{RT} \qquad 2.26$$

Steady-state anisotropy (r_{ss}) for the single-exponential intensity decay can be obtained by Perrin equation.

$$r_{ss} = \frac{r_0}{1 + \frac{\tau}{\theta}} \qquad 2.27$$

Here r_0 denotes initial anisotropy, θ is the rotational correlation time of the molecule and τ represents the lifetime of the fluorophore.

Initial anisotropy (r_0) of a molecule can be written as

$$r_0 = \frac{2}{5}\left(\frac{3 < cos^2 \beta > -1}{2}\right) \qquad 2.28$$

The value 2/5 comes from the principle of photo-selection for randomly orientated fluorophore in solution. The angle between emission and absorption dipoles is given by β. r_0 is also known as fundamental anisotropy.

If a sample consists of different emitting species, average anisotropy $\bar{r}$ can be defined as

$$\bar{r} = \sum_i f_i \, r_i \qquad 2.29$$

Here r_i denotes anisotropy of each species and f_i express fractional amplitude of each species and $\sum_i f_i = 1$.

For a system containing multiple fluorophore or single fluorophore showing multiple rotational times, anisotropy decay at time t will be multi-exponential and can be expressed as

$$r(t) = r_0 \sum_j \beta_j \exp(-t/\theta_j), j<3 \qquad 2.30$$

Here θ_j is the correlation time of individual species and β_j is the fractional amplitude of each correlation time.

Time-resolved anisotropy is frequently used to study protein-protein and protein-nucleic acid interaction (Hill and Royer, 1997; Rusinova et al., 2002). Stryer and co-workers analyzed the rotational motion in the antibody labeled with dansyl probe. Dansyl labeled antibody exhibit two type of rotational motion: 1) Slow rotation motion (global motion) about 168 ns arising due to the tumbling of whole antibody and 2) Fast rotation motion (segmental motion) about 33 ns arising due to the movement in F_{ab} (Yguerabide et al., 1970).

Baird and co-workers studied the anisotropy decays of IgE in its free state and when it is bound to a receptor in membrane. When IgE is bound to the receptor (Fc_{ε}, type I), it shows a long correlation time around 438 ns, while IgE free in solution, exhibits two, 48 and 125 ns correlation times (Holowka et al., 1990).

Molecules may display multiple rotational correlation times because of non-spherical symmetry. Rate of rotation will be different around each axis for non-spherical molecules. For non-spherical molecules two types of ellipsoid of revolution 1) prolate and 2) oblate are used to explain anisotropy decays. Theoretically, non-spherical molecules can display five correlation times but in practice only three correlation times are possible because the magnitude of other two correlation time will be very close and will not resolved easily (Small and Isenberg, 1977).

2.1.2.2.5. Time-resolved intensity decay analysis

Time-resolved fluorescence data collected through TCSPC are quite complex and simple graphical methods cannot be used to evaluate the data. Intensity decay curve contains three curves; Instrument Response Function (IRF) [$L(t_k)$], measured intensity decay $N(t_k)$ and calculated (fitted) intensity decay $Nc(t_k)$. The measured intensity decay profile $N(t_k)$ is a convolution of the IRF $L(t_k)$ with actual intensity decay of the sample, $Nc(t_k)$. These functions are in term of discrete time (t_k) because the counted photons are collected into channel each with known width (Δt) and time t_k.

The instrument response function (IRF) [L(*tk)*] also known as the lamp function is an instrument response to sample with zero fluorescence lifetime. This decay shows the shortest time that can be detected by TCSPC. The IRF can be collected by diluted non-fluorescence scattering solution like colloidal silica or $Mg(OH)_2$ in absence of any emission filter. The IRF used for our experiments was finely granulated chalk powder. The intensity decay of sample measured by TCSPC is called as measured intensity decay $N(t_k)$. The measured intensity decay $N(t_k)$ contains significant contribution from IRF. This has to be analyzed to determine the actual lifetime of the sample. The mathematical approach used to extract the intensity decay parameters from measured intensity decay is known as iterative reconvolution.

The impulse response is observed if IRF is a δ-function. The concept of convolution can be represented as

$$I_k(t) = L(t_k)I(t - t_k)\Delta t \ (t > t_k) \qquad 2.31$$

The impulse response started at $t=t_k$ because of this term $(t-t_k)$ appears. It is considered that there is no emission prior to excitation (at $t=t_k$). The term Δt describes the width of the channel in which the counted photons are collected.

The measured intensity decay $N(t_k)$ can be defined as the sum of impulse response formed by all the individual δ-function excitation pulses appearing until t_k.

$$N(t_k) = \sum_{t=0}^{t=t_k} L(t_k)\, I(t - t_k)\Delta t \qquad 2.32$$

Above equation can be represented as an integral, for small value of Δt.

$$N(t_k) = \int_0^t L(t')\, I(t - t')dt' \qquad 2.33$$

According to the above expression the measured intensity at time t is a sum of the intensities presumed for all the δ-function excitation pulses that appears until time t. The new intensity decays are generated in the sample until there is nonzero intensity in $L(t_k)$.

Several methods like non-linear least square analysis (NLLS), the maximum entropy method, the method-of-moments, Laplace transformation and phase-plane method etc.

are used to analyze TCSPC data. The Marquardt's algorithm based non-linear least square analysis (NLLS) is broadly used to analyze the complex intensity decay.

A least square analysis starts with a model that is expected to be the best representation of the data.

The guess values of I_0 and τ is assumed for a mono-exponential decay $I(t) = I_0\, exp^{-t/\tau}$. To know the value of calculated or fitted decay, $Nc(t_k)$ these guess values are convoluted with the IRF. Further, this calculated decay $Nc(t_k)$ is compared with experimentally resolved decay $N(t_k)$. The minimized value of goodness of fit parameter χ^2 is calculated to judge the goodness of fit. The χ^2 can be expressed as

$$\chi^2 = \sum_{k=1}^{n} \frac{1}{\sigma_k^2} [N(t_k) - N_c(t_k)]^2 \qquad 2.34$$

$$= \sum_{k=1}^{n} \frac{[N(t_k) - N_c(t_k)]^2}{N(t_k)} \qquad 2.35$$

The χ^2 is not convenient to interpret as it depends upon numbers of data points. Hence another parameter is used known as reduced χ^2 (χ_R^2) and can be represent as

$$\chi_R^2 = \frac{\chi^2}{n-p} = \frac{\chi^2}{v} \qquad 2.36$$

Here, n represents the number of data points, p denotes the number of floating parameters, and v is the number of degree of freedom.

If calculated and experimentally obtained decay data are closely matched then value of χ_R^2 will be close to unity. If the data does not fit properly then value of χ_R^2 will be higher than 1. In this case new set of parameters are selected by modifying the older parameters. Until the value of χ_R^2 does not reach close to unity, process of matching and convolution are repeated iteratively. This operation is termed as iterative reconvolution.

Another parameter to judge the quality of fit is known as deviation plots or distribution of residuals.

The difference between measured and fitted function yields the deviation (D_k). Deviation plot is obtained by plotting standard deviation at each data point against time. The deviation (D_k) can be expressed as

$$D_k = \frac{I(t_k) - I_c(t_k)}{\sqrt{I(t_k)}} \qquad 2.37$$

Goodness of the fit can be judged by checking the randomness of residual distribution. An improper fitting of decay is indicated by a distinct pattern in the distribution of residual.

For MEM analysis, the decays were fitted to 100 exponentials with lifetime values between 0.01-20 ns, distributed uniformly in the logarithmic scale. Each iteration is attempted to achieve a distribution of lifetimes by minimizing the χ^2 and maximizing the entropy, S (known as Shannon-Jaynes Entropy) given by equation 2.38 and 2.39, respectively. The goodness of fit was determined by the value of obtained reduced chi square (χ^2_R) value a well as the increased entropy value. At least three such fits were performed for each sample and the fit with best χ^2_R values were selected.

$$S = \sum_i -p_i \log p_i \qquad 2.38$$

Where, p_i is the probability (amplitude) for the ith lifetime. All lifetimes in the distribution are assigned an equal amplitude (0.001) in the beginning. Among many possible distributions of p_i for a given value of χ^2, MEM gives the distribution for which entropy, S is maximum. Here, χ^2 is given by:

$$\chi^2 = \frac{\left(\frac{1}{m}\right)\sum_{i=1}^{m}[F_C(t_i)-F_e(t_i)]^2}{F_e(t_i)} \qquad 2.39$$

Where, m is the number of data points and $F_C(t_i)$ and $F_e(t_i)$ are the calculated and experimental fluorescence intensity, respectively at time t_i.

2.1.2.2.6. Time-resolved Anisotropy decay analysis

The time-resolved anisotropy decay data was analyzed on the basis of equations given below:

$$I_{\parallel}(t) = \frac{1}{3} I(t)[1 + 2r(t)] \qquad 2.40$$

$$I_{\perp}(t) = \frac{1}{3} I(t)[1 - r(t)] \qquad 2.41$$

$$r(t) = r_0 \sum_i g_i exp(-t/\theta_i) \quad i = 1 \, or \, 2 \qquad 2.42$$

Here, r_0 denotes the initial anisotropy, amplitude of ith rotational correlation time θ_i is denoted as g_i and $\sum_i g_i = 1$. $I_{\parallel}(t)$ and $I_{\perp}(t)$ represents the sample intensity decay measured with the polarizers in the parallel and perpendicular orientations, respectively.

If biomolecules shows the biexponential $r(t)$ then θ_{fast} and θ_{slow} reveals the information of the local motion of the labeled probe and global motion (overall tumbling) of the biomolecules, respectively. To extract the best value for g_i and θ_i, a non-linear least square fitting model based on Marquardt algorithm (Bevington and Robinson, 1992) is executed. The measured instrument response function (IRF) is used to analyze the $I_{\parallel}$ and $I_{\perp}$.

The values collected in the above fitting analysis were used to yield the calculated steady-state anisotropy r_{ss} (Swaminathan et al., 1994).

$$r_{ss} = \frac{r_0 \sum_i \sum_j \alpha_i \, g_i \left(\frac{1}{\tau_i} + \frac{1}{\theta_i}\right)^{-1}}{\sum_i \alpha_i \tau_i} \qquad 2.43$$

Further, the calculated steady-state anisotropy values were compared with the steady-state anisotropy values measured independently by using spectrofluorometer, to cross verify the fitted values of θ_i, g_i and r_0.

2.1.2.2.7. Förster Resonance Energy Transfer (FRET)

Resonance Energy Transfer (RET) is an important process which occurs in the excited state. RET occurs when there is an overlap between emission spectrum of a fluorophore (donor) and the absorption spectrum of a chromophore (acceptor). FRET is an important tool, utilized for studying the protein-protein interaction, protein folding and protein aggregation. FRET is an electrodynamic process; involve transfer of excited state energy from excited state donor (D) to ground state acceptor (A) molecule. Transfer of energy occurs via non-radiative dipole-dipole interaction between the donor and acceptor. Rate of energy transfer depends upon; a) spectral overlap between donor emission spectrum and the acceptor absorption spectrum, b) orientation of transition dipoles between the donor and acceptor, c) the quantum yield of donor and d) distance between donor and acceptor molecules. Efficiency of energy transfer is very sensitive to the Förster radius (R_0) ranging from 10 to 100 Å. Förster radius can be defined as the distance at which efficiency of energy transfer is 50%. The value of R_0 depends upon the spectral properties of donor and acceptor fluorophore. Förster radius (R_0) can be expressed as

$$R_0^6 = 8.79 \times 10^{23} \left[\kappa^2 n^{-4} Q_D J(\lambda)\right] (\text{Å}) \qquad 2.44$$

Here, κ^2 denotes dipole orientation factor, η represents refractive index, Q_D denotes quantum yield of donor in absence of acceptor and $J(\lambda)$ denotes spectral overlap integral. $J(\lambda)$ can be expressed as

$$J(\lambda) = \int_0^\infty F_D(\lambda)\, \varepsilon_A(\lambda)\, \lambda^4 d\lambda\, M^{-1} cm^3 \qquad 2.45$$

Here, ε_A represents the extinction coefficient of acceptor and F_D denotes fluorescence emission intensity of donor as a fraction of the total integrated intensity.

The rate of energy transfer $k_T(r)$ is expressed as

$$k_T(r) = \left(\frac{1}{\tau_D}\right)\left(\frac{R_0}{r}\right)^6 \qquad 2.46$$

Here, r is the distance between the donor and the acceptor, R_0 denotes the Förster radius and τ_D is the life time of the donor in absence of acceptor.

For a single donor-acceptor pair at a fixed distance, the efficiency of energy transfer can be expressed as

$$E = \frac{R_0^6}{R_0^6 + r^6} \qquad 2.47$$

The transfer efficiency is calculated by taking the relative fluorescence intensity of the donor, in the absence of acceptor (F_D) and presence of acceptor (F_{DA}) and expressed as

$$E = 1 - \left(\frac{F_{DA}}{F_D}\right) \qquad 2.48$$

The transfer efficiency can also be calculated from the lifetimes of donor, in absence and presence of acceptor.

$$E = 1 - \left(\frac{\tau_{DA}}{\tau_D}\right) \qquad 2.49$$

Here, τ_{DA} is the lifetime of donor in presence of acceptor and τ_D is the lifetime of donor in absence of donor.

2.1.3. Circular Dichroism

Circular Dichroism (CD) is a powerful technique to detect the changes in secondary structure of proteins and nucleic acids in solution. To display the CD signal, chromophore should have chirality or present in the asymmetric environment.

Parameter collected in the CD is the net difference in absorbance between right and left handed circularly polarized component of plane polarized light. Difference in absorbance can be represent as

$$\Delta A = A_L - A_R \qquad 2.50$$

Where A_L and A_R denote absorbance for left and right-handed circularly polarized light, respectively.

As the shape of CD and absorption bands are usually similar, CD spectrum can be predicted from absorption spectrum by calculating the area under each CD band and noting sign. These quantities together are known as the rotational strength (R_{0a}) for state 0 to state a transition. The rotational strength is equivalent to the dipole strength used as a measure of absorption intensity.

$$R_{0a} = (3hc/8\pi^3 N_0) \int \{[\theta(\lambda)]/\lambda\}\, d\lambda \qquad 2.51$$

Where, h denotes Planck’s constant and c is the speed of light.

Rotational strength can be determined from principles of quantum mechanical and knowledge of ground (ψ_0) and excited (ψ_a) states wavefunction of a chiral molecule.

$$R_{0a} = imaginary(\langle\psi_0|\tilde{\mu}|\psi_a\rangle \cdot \langle\psi_a|\tilde{m}|\psi_0\rangle) \qquad 2.52$$

Where, $\tilde{\mu}$ denotes an electric dipole operator and $\tilde{m}$ magnetic dipole operator.

When plane polarized light pass through optically active sample, one of the circular components is absorbed more than the other. Instead of plane polarized light, elliptically polarized light is obtained by the recombination of unequally absorbed components. Practically, CD instrument separately detect the two-component instead of recombined them. Relationship between Circular Dichroism and ellipticity can be expressed as

$$[\theta] = 3300\ \Delta\mathcal{E} \qquad 2.53$$

Here, [θ] represent the molar ellipticity and most of circular dichroism measurements are based on ellipticity determinations.

The far-UV CD spectra basically reveal the information about the secondary structure content in the protein like α-helix, β-sheet, β-turn and statistical coil. In the region of 240 to 190 nm, light is majorly absorbed by the peptide bond leading to a strong π to π^* (~ 190 nm) transition and weak but broad n to π^* transition (210-220 nm).

Near-UV CD spectra reveal information about the tertiary structure of the protein. CD spectra in the region of 250 to 290 nm basically arises by absorption of aromatic amino acid side chain such as Phe, Tyr and Trp. Folding of polypeptide chain can locate side chain of aromatic amino acids in the chiral environment. This enables the CD spectra to provide a characteristic fingerprint of folded protein.

2.2. Materials

Acrylamide (A3553); Acetonitrile (34998); Ampicillin (A8351); Agarose(A9539); EtBr(E7637); Isopropyl β-D-1-thiogalactopyranoside (I5502); Phenylmethanesulfonylfluoride (P7626); Nickel Affinity gel (H0537); Triton X-100 (T9284); β Mercaptoethanol (63689); Glycine (50046); Sinapic acid (85429); $CaCl_2$ (C8106); Gel loading buffer (G2526); Bovine serum albumin (A3059); Trifluoroacetic acid (T6508); 2,2,2-Trifluoroethanol (T8132); Blue dextran (D4772); Guanidine Hydrochloride (G3272); 8-Anilinonaphthalene-1-sulfonic acid (A3125);; N-acetyl tryptophan amide (A6501); 2,2′-Dithiobis(5-nitropyridine) (158194); L-Cysteine (168149); Ammonium persulfate (A3678); N,N,N′,N′-Tetramethylethylenediamine (A7024); Chorismic acid (C1761) and Hen Egg-white Lysozyme (L6876) protein were purchased from Sigma Aldrich Chemicals Pvt. Limited, Bengaluru, India.

The chemicals for synthesis of the Transition State Analogue (TSA) (8-HYDROXY-2-OXA-BICYCLO[3.3.1]NON-6-ENE-3,5-DICARBOXYLIC ACID) namely 1,3-Butadiene (295035); Butylated hydroxytoluene (W218405); Dimethyl Itaconate (109533); Nafion NR50 (309389); TMSBr (Trimethylsilyl bromide) (194409) and Zinc iodide (223883) were also purchased from Sigma Aldrich Chemicals Pvt. Limited, Bengaluru, India.

IAEDANS (1,5-IAEDANS, 5-((((2-Iodoacetyl)amino)ethyl)amino)Naphthalene-1-Sulfonic Acid) (I14) was purchased from Invitrogen (Carlsbad, California, USA). Luria Broth (M1245); Luria Agar (M557); Terrific Growth Medium (G004); Magnesium chloride (TC186) and Imidazole (GRM1864) were procured from HiMedia Laboratories, India. The PD-10 desalting columns (17-0851-01) and Ni Sepharose His-Trap FF affinity columns (17525501) were purchased from GE Healthcare. The Tris (Hydroxy methyl amino methane) (93315); Citric acid monohydrate (100244); Sodium acetate (17952); di-Sodium hydrogen phosphate (17549); Sodium chloride (40731); Dimethylformamide (17754); Dimethyl sulfoxide (61857); SDS (184190); Potassium dihydrogen phosphate (104873); Potassium chloride (91635) were purchased from Merck India Limited. All chemicals used were analytical grade with purity ≥98%.

MassRuler DNA Ladder Mix (SM0403) from Thermo Scientific; Ultra Low Range Molecular Weight Marker (M.W. 1,060–26,600) (M3546) from Sigma Aldrich Chemicals Pvt. Limited, Bengaluru, India.

2.3. Experimental methods

2.3.1. Site directed mutagenesis

In accordance with our mutation plans, primers were designed using the Agilent QuikChange Primer designing software and the T_m values obtained were cross-checked using the formula

$$T_m = 81.5 + 0.41(\%GC) - (675/N) - \%mismatch \qquad 2.54$$

Where N is the primer length in bases and values for **%GC** and **% mismatch** are whole numbers. All the primers were purchased in lyophilized powder form from Integrated DNA Technologies, Inc. The stock solutions were prepared in 20 mM Tris buffer pH 8.0 and stored in -20^{0}C.

The parent DNA plasmid (mMjCM) used as the template for mutagenesis was generously gifted by Prof. Donald Hilvert's lab in ETH Zurich, Switzerland.

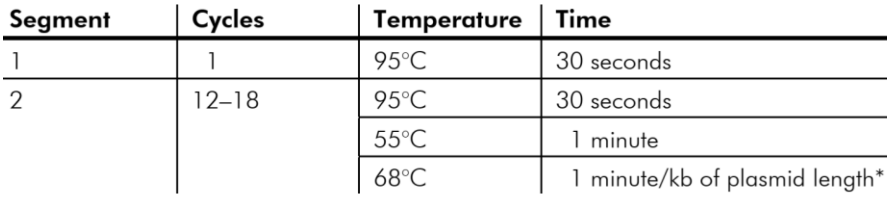
The mutations were carried out using the QuikChange II Site-directed Mutagenesis Kit (200524) with minor modifications. The sample composition used mutations for PCR-like reaction is listed below in **Table 2**.

Table 2. Reaction mixture for PCR-like thermal cycling

Control reaction	Sample reaction(s)
5 μl of 10× reaction buffer 2 μl (10 ng) of pWhitescript 4.5-kb control plasmid (5 ng/μl) 1.25 μl (125 ng) of oligonucleotide control primer #1 [34-mer (100 ng/μl)] 1.25 μl (125 ng) of oligonucleotide control primer #2 [34-mer (100 ng/μl)] 1 μl of dNTP mix 38.5 μl ddH2O (to bring the final reaction volume to 50 μl)	5 μl of 10× reaction buffer X μl (5–50 ng) of dsDNA template X μl (125 ng) of oligonucleotide primer #1 X μl (125 ng) of oligonucleotide primer #2 1 μl of dNTP mix ddH_2O to a final volume of 50 μl

Then1 μl of PfuUltra HF DNA polymerase (2.5 U/μl) was added to both control and sample reactions.

The parameters used for the PCR-like temperature cycling are as follows:



Segment	Cycles	Temperature	Time
1	1	95°C	30 seconds
2	12–18	95°C	30 seconds
		55°C	1 minute
		68°C	1 minute/kb of plasmid length*

Following the PCR-like reactions using a thermal cycler (make: Applied Biosystems) were placed on ice for 2 minutes to cool the reaction to ≤37^{0}C. For the final step of this mutagenesis process, 1 µl of the *Dpn* I restriction enzyme (10 U/µl) is added directly to each amplification using a small, pointed pipet tip. Each reaction mixture is gently and thoroughly mixed by pipetting the solution up and down several times. The sample mixtures are then spun in a microcentrifuge for 1 minute and immediately incubate each reaction at 37^0C for 1 hour to digest the parental (i.e., the nonmutated) supercoiled dsDNA. The daughter DNA plasmids containing the desired mutations are then transformed into competent cells for further amplification and isolation of colonies. The positive mutant colonies obtained after transformation were confirmed by DNA sequencing through Scigenom pvt. Ltd.

The Dpn I digested DNA and newly formed daughter plasmids were visualized using Agarose Gel Electrophoresis.

2.3.2. Competent cell preparation and transformation of plasmid DNA

E. coli DH5α and BL-21 (DE3) competent cells were prepared for plasmid DNA amplification/isolation and protein expression respectively. A single bacterial colony was carefully picked up from a Luria Agar plate and inoculated into 5 mL of LB medium. The culture was allowed to grow in a shaking incubator at 37 ^{o}C and 180 rpm for overnight. 500 µL (1% of the final culture volume) overnight grown culture was used to inoculate in 50 mL of LB medium. This culture was allowed to grow at 37 ^{o}C at 180 rpm till the A_{600} of the culture reached approximately 0.4. Subsequently, culture was incubated at 4 °C for 10 minutes and cells were collected by centrifugation at 3000 rpm for 10 minutes at 4 °C. After discarding the supernatant, cells were resuspended gently by pipetting or inverting in 15 mL of ice cold 80 mM $MgCl_2$ and 20 mM $CaCl_2$ solution. Cells were centrifuged again at 3000 rpm for 10 minutes at 4 ^{o}C. The supernatant was discarded and collected pellet was re-suspended in 900 µL of 0.1 M $CaCl_2$ and 100 µL of glycerol. The cells were kept as aliquots of 100 µL in microcentrifuge tubes. The competent cells were stored at -80 ^{o}C until further use. All the materials were autoclaved and all the steps were carried out in laminar air flow to maintain proper sterile conditions.

The mutant PCR-like products were transformed into the DH5α competent cells using heat shock at 42 ^{o}C for 60 secs followed by immediate incubation in ice for 10 mins. 800 µL of sterile LB media (without any antibiotic) was then added to the cells and was incubated in a shaking incubator at 37 °C and 180 rpm for 45 minutes. The cells were then harvested by centrifugation at 6,000 rpm for 10 minutes. After discarding supernatant, the collected pellet was resuspended in 100 µL of fresh LB media and plated on Luria Agar plates containing Ampicillin as the antibiotic for selection (mMjCM plasmid contains Amp^r). A range of 10-20 positive colonies were obtained for the mutant plasmids. The mCM wild-type plasmid and the mutant plasmids were also transformed into BL21 (DE3) bacterial stains for expression of the individual proteins and their subsequent purifications.

All plasmids were isolated from transformed DH5α cells using the Sigma GenElute™ Plasmid Miniprep Kit (PLN70) and stored in Tris buffer pH 8.0 at -20 ^{o}C for PCR-like amplifications and DNA sequencing.

2.3.3. Expression and purification of mCM wild-type and mutants

A single colony of transformed *E. coli* BL21 (DE3) cells was picked up from a LB agar plate and was used for over expression of all mCM variants. The cultures were grown in LB media, containing 100 µg/mL ampicillin (at 37^{o}C, 180 rpm) in a shaking incubator. At A_{600} of 0.6-0.7, cultures were induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) at 37 ^{o}C for 3 hours. Induced cells were centrifuged at 5000 rpm for 15 minutes at 4 ^{o}C to collect the pellet. The collected pellets were resuspended in ice cold lysis buffer (20 Tris.HCl (pH 7.4); 150 mM NaCl; 0.5% Triton-X; 1% glycerol; 1mM PMSF) and kept at 4^{o}C. Subsequently, resuspended cells were disrupted by sonication at 33% power output with pulse cycles of 4 seconds on and 10 seconds off for 20-30 minutes (depending on the lysate volume) at 4 ^{o}C. To remove the cell debris, cell lysates were centrifuged at 12000 rpm for 30 minutes at 4 ^{o}C. Supernatants were filtered with a 0.45 µm filter (make: Millipore). The filtered supernatants were then allowed to bind to a pre-equilibrated (binding buffer 20mM Tris.HCl (pH 7.4); 150 mM NaCl; 20 mM Imidazole) Ni Sephrose HisTrap column (5ml) for 1 L starting culture. The supernatant was circulated through the column for 4-5 hours using a peristaltic pump (make: GE Healthcare) operating at a speed of 0.5 ml/min. Following protein binding, the Ni Sepharose column was washed with a wash buffer (20 mM Tris.HCl (pH 7.4); 150 mM NaCl; 75 mM Imidazole) for approximately 2 hours (5-8 column volumes). The bound protein was eluted with an elution buffer (20 mM Tris.HCl (pH 7.4); 150 mM NaCl; 350 mM Imidazole).

Purity of all the mCM variants were checked by SDS polyacrylamide gel electrophoresis (15% Acrylamide). The fractions containing pure proteins were pooled and dialyzed against dialysis buffer 1 (PBS pH 7.4; 150 mM NaCl; 100 mM EDTA) to slowly chelate any leaked Ni^{2+} ions (from the column) and remove imidazole. This was followed by dialysis buffer 2 (PBS pH 7.4; 150 mM NaCl) to remove EDTA and any left-over contaminants. Purified proteins were further analyzed by mass spectrometry to ascertain the identity and purity of purified proteins.

2.3.4. SDS PAGE electrophoresis

Reducing SDS-PAGE was performed according to the protocol reported in Sambrook and Russel, Molecular Cloning-A laboratory manual (Sambrook et al., 1989) to check the purity of purified protein. Resolving gel (15% acrylamide) was used for running all the purified protein samples. Composition of SDS-PAGE is given in appendix. Protein sample was mixed with the gel loading buffer in 4:1 ratio and subsequently allowed to boil at 95 ^{o}C for 5 minutes. Samples were separated on vertical Mini-PROTEAN^{R} Tetra Electrophoresis System (Make: Bio-Rad). All protein samples were run at 80 Volts for 2 hours. After running the samples, gels were removed from electrophoresis apparatus and stained with colloidal Coomassie stain for 4 hours. After proper staining, the gels were placed in de-staining solution (see appendix) for overnight to allow the de-staining of gel. Standard molecular markers (Sigma Aldrich ultra-low range molecular marker, M.W. 1,060–26,600) were used to determine the approximate molecular weight of protein in the gel.

2.3.5. Mass spectrometry

The mass of mCM wildtype (Wt mCM) and the mutants were determined by using a MALDI-TOF mass spectrometer (Bruker Daltonics, Germany). Matrix was prepared in TA-30 solvent by saturating solution of sinapic acid. The saturated solution of sinapinic acid was then sonicated for 30 mins and centrifuged at 12000 rpm at room temperature (25 ^{o}C) for 15 mins. The supernatant obtained was then mixed with protein samples in 2:1 ratios and spotted on the target plate. The spotted samples are allowed to dry at room temperature for 2 hours. All data acquisitions were done in linear mode with flexControl and analysis done with flexAnalysis software from Brunker Daltonics.

2.3.6 Synthesis of TSA

The Transition State Analogue (TSA), (8-hydroxy-2-oxa-bicyclo[3.3.1]non-6-ene-3,5-dicarboxylic acid), was synthesized by Dr. P. Gangadhar in Prof. P.Srihari’s lab (IICT

Hyderabad, India). The synthesis process was accomplished by using Bartlett's published protocol (Bartlett and Johnson, 1985).

2.3.7. UV-Visible absorption spectra

The absorption spectrum of all mCM variants were recorded using double beam Lambda-25 UV-Vis Spectrophotometer (Perkin Elmer, USA) at room temperature (25 ^{o}C). UV transmissible quartz cuvette of 10 mm path length (Hellma; Z600210) was used for all absorption measurements. The absorption spectra were acquired with multiple scans (3) between 250 nm and 700 nm (1 nm bandwidth; 480 nm/minute scan speed) and averaged subsequently. All spectra shown are averages of at least two independent measurements. Blanks (water/ PBS buffer) were identical to the sample in all respects with the exception that they contained no protein.

Extinction coefficients (ε) at $\lambda_{280\ nm}$ 15595 $M^{-1}cm^{-1}$ for Wt mCM, 10095 $M^{-1}cm^{-1}$ for Mutant1 and Mutant3, 9975 $M^{-1}cm^{-1}$ for Mutant2 were used for the estimation of protein concentration using the Beer-Lambert equation mentioned earlier in this chapter.

2.3.8. mCM labelling with dansyl probe

Labeling of single cysteine in mCM Wt, Mutant 1 and Mutant 3 with the IAEDANS (1,5-IAEDANS, 5-((((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulphonic acid) was carried out as recommended by Molecular Probe with slight modifications (Wiederschain, 2011). 100 mM stock of IAEDANS was prepared by dissolving the probe into the DMF. 100 µL of 100 mM IAEDANS was added (final concentration of IAEDANS was 10 mM) slowly to 900 µL of 100 µM mCM (20 mM Tris.HCl, pH 9; 150 mM NaCl) to make a total reaction volume of 1 ml. This reaction mixture was incubated at room temperature (25 °C) for three hours with continuous stirring at 150 rpm. On completion of the reaction, the free dye was removed from the dansyl labeled protein by using pre-equilibrated PD-10 desalting column with 20 mM Tris.HCl (pH 9); 150 mM NaCl. The labeled protein was eluted from the column in PBS buffer pH 7.4. Concentration of protein and dye was measured by using absorbance at 280 and 337 nm, respectively. The molar extinction coefficient of conjugated probe with protein at 337 nm is 6100 $M^{-1}cm^{-1}$ (Hudson and Weber, 1973).

2.3.9. Steady state fluorescence and anisotropy measurements

All steady state fluorescence and anisotropy measurements were performed using Fluoromax-4 Spectrofluorometer (Jobin-Yvon Horiba Inc., USA) equipped with motorized polarizers. For all tryptophan steady state fluorescence intensity measurements of mCM (mCM Wt and all mutants), samples were excited at 295 nm (slit width = 2 nm) and emission was collected between 310 nm to 500 nm (slit width = 5 nm). For acquiring tryptophan steady state anisotropy of mCM, all samples were excited at fixed 295 nm (slit width = 2 nm) and emission was collected at fixed 345 nm (slit width = 5 nm).

For acquiring steady state fluorescence intensity of dansyl labeled with cysteine of mCM, all samples were excited at 340 nm (slit width = 2 nm) and emission was collected between 355 nm to 650 nm (slit width = 5 nm). For collecting steady state anisotropy of dansyl labeled with mCM, all samples were excited at fixed 340 nm (slit width = 2 nm) and emission was collected at fixed 505 nm (slit width = 5 nm). All steady state anisotropy data were G-factor corrected and each anisotropy measurement was an average of 10 independent measurements.

The background intensity from buffers and Raman scatter were insignificant (<5%) in comparison with fluorescence intensity of samples under the identical conditions. However, these were subtracted from emission spectra of sample. All steady state fluorescence measurements were repeated at least three times at room temperature (25 ^{o}C) in 200 μL quartz cuvette with a 10 mm pathlength.

2.3.10. ANS binding assay

To study the binding of 8-Anilinonaphthalene-1-sulfonic acid (ANS) with mCM, stock of ANS (1mg/mL) was prepared in deionized water. Stock concentration of ANS was confirmed by using extinction coefficient of 4,950 $M^{-1}cm^{-1}$ at 350 nm. 20 μM of mCM (dissolved in PBS pH 7.4) was mixed with 40 μM of ANS.

To study the effect of TSA on ability of ANS to bind the ligand-bound mCM (all variants), 40 μM of TSA was added to 20 μM of mCM and incubated for an hour before addition of 40 μM of ANS to the sample mixture.

Binding of ANS with mCM was monitored by measuring the steady state fluorescence intensity of ANS. All samples were excited at 380 nm (slit width = 2 nm) and emission was collected between 400 nm to 700 nm (slit width 10 nm). The background intensity from buffer and Raman scatter were negligible (<5%) in comparison with fluorescence intensity of sample under the identical condition. However, these were subtracted from emission spectra of sample. All ANS fluorescence measurements were done in triplicate at room temperature (25^{o}C).

2.3.11. Time-resolved fluorescence and anisotropy measurements

All time resolved fluorescence intensity decay and anisotropy decay measurements were done by time-correlated single photon counting (TCSPC) instrument, using the ‘Delta-Pro^{TM}’ equipped with motorized polarizer, supplied by Horiba Scientific, UK. To measure the fluorescence lifetime and anisotropy decay of tryptophan or dansyl probe in mCM under various conditions, all samples were excited with pulsed light source, 295 nm DeltaDiode^{TM} having about 0.81 ns Instrument response function (IRF) FWHM and a 20MHz repetition rate. Emission of tryptophan was collected with 320 nm longpass filter (WG320) to block the excitation light. Total peak counts were collected up to 15,000 counts and fluorescence intensity decay was collected in 2202 channel (in 0 to 61.66 ns time range) with temporal resolution of 0.028 ns/ channel.

To measure the fluorescence intensity and anisotropy decay of dansyl, labeled with cysteine of mCM (Cys69/6), all samples were excited with pulsed light source, 340 nm DeltaDiode^{TM} having about 0.80 ns IRF FWHM and a repetition rate of 20MHz. Emission of dansyl was measured with 370 nm longpass filter (KV370) to cut the excitation light. Total peak counts were measured up to 20,000 and fluorescence intensity decay was collected in 3995 channel (in 0 to 223.72 ns time range) with temporal resolution of 0.056 ns/channel. Emission of all samples were measured at magic angle position of 54.7° with emission polarizer. IRF was collected using scattering solution of colloidal silica, without emission filter for each of the sample. All lifetime and anisotropy measurements were done at room temperature (25 ^{o}C) and at least three independent measurements were recorded.

All time-resolved anisotropy experiments was performed with the emission polarizer oriented perpendicular ($I_{\perp}$) and parallel ($I_{\parallel}$) to the excitation polarizer. DAS6 v6.8 decay analysis software provided by *HORIBA Scientific* was employed to obtain lifetime of tryptophan from intensity decay data for all conditions. Lifetime of dansyl labeled with cysteine of mCM was obtained using *exp 2* and *exp3* decay analysis software (kindly provided by Dr. N. Periasamy, Tata Institute of Fundamental Research (TIFR), Mumbai, India) for all conditions. Rotational correlation times of dansyl (for all conditions) were extracted from *tani* software kindly provided by Dr. N. Periasamy, TIFR. The goodness of fit was evaluated by analysing the reduced χ^2 and randomness of residuals.

2.3.12. Intramolecular Förster Resonance Energy Transfer (FRET) measurements

The intramolecular distance between Trp24/26 and dansyl labeled Cys69/6 in the mCM proteins were measured in absence and presence of varying concentrations of the ligand (TSA). To calculate the distance between tryptophan and cysteine in all the cases, fluorescence lifetime of tryptophan in absence of the acceptor (τ_D), dansyl, and in presence of the acceptor (τ_{DA}) was monitored. The fluorescence lifetime of tryptophan was collected by TCSPC. The tryptophan was excited with pulsed light source, 295 nm DeltaDiode^{TM} and subsequently emission of tryptophan was collected with 330-355 nm bandpass filter (ET340/40X) supplied by *CHROMA*, to avoid collecting emission light coming from the dansyl probe. Total peak counts were measured up to 15,000 and fluorescence intensity decay was collected in 2202 channel (in 0 to 61.66 ns time range) with temporal resolution of 0.028 ns/channel. All lifetime measurements were done at room temperature and for each samples at least three independent lifetime measurements were recorded.

Förster distance (R_0) 22 Å, for tryptophan-IAEDANS pair was used (Wu and Brand, 1994). FRET efficiency (E) was calculated from mean lifetime of donor alone (τ_D) and in presence of acceptor (τ_{DA}) by using equation 2.46.

For FRET analysis, the K value used was 0.816 (κ^2 = 2/3) and value for η used was 1.00 cP (Khrenova, Topol et al. 2015).

2.3.13. Circular Dichroism analysis:

CD spectra of mCM were recorded on a CD spectrometer (Make: Jasco, Model J-1500, Jasco Inc., Maryland, USA), equipped with a Peltier thermoregulator.

To measure the CD spectra of mCM (all variants), protein samples were prepared by dissolving 10 μM of mCM in 10 mM phosphate buffer supplemented with 10 mM NaF (pH 7.4). All CD spectra were collected from 190 to 260 nm, at a scan speed of 100 nm/minute, bandwidth of 2 nm and data pitch of 0.1 nm. All spectra were collected using 2 mm quartz cuvette (Starna Scientific Ltd). The contribution of deionized water and respective buffers were subtracted from the acquired spectra during data acquisition. For each protein sample at least three independent scans were carried out. CD spectra from 190 to 260 nm were analyzed using K2D3 server. A rough estimate of the α-helical content and its variation across the mCM proteins could be was made using this analysis.

Chemical denaturation of Wt mCM and its mutants were performed by incubating protein samples in increasing concentration of GdnHCl (0-6 M). Changes in the secondary structure were monitored by following the changes in ellipticity at 222 nm at 25 ^{0}C. Data points obtained were fitted by non-linear regression analysis (dose-response curve) using OriginPro 8 software.

For thermal denaturation, change in ellipticity of Wt mCM and the mutants were continuously monitored at 222 nm wavelength in the temperature range from 25 to 90 ^{0}C and back from 90 to 25 ^{0}C with a heating/cooling rate of 1 ^{0}C/min. CD spectral scans from 260 nm to 190 nm at specific temperature intervals and their reverse scans were also recorded during the experiment.

The data output from the CD spectrometer (in mdeg) was converted to MRE (deg .cm^2. dmol^{-1}) using the formula below:

$$MRE = Y \ m\deg \times \frac{1\deg}{10^3 m\deg} \times \frac{liter}{C\ mol} \times \frac{10^3 cm^3}{liter} \times \frac{1}{0.1\ or\ 1\ cm\ path} \times \frac{mol}{10^1 dmol} \times \frac{1}{n_{res}} \qquad 2.55$$

where Y is the signal measured, C is the concentration in molar units, and n_{res} is the length in residues of the protein.

We used a 0.2 cm pathlength cuvette, hence calculations were done accordingly.

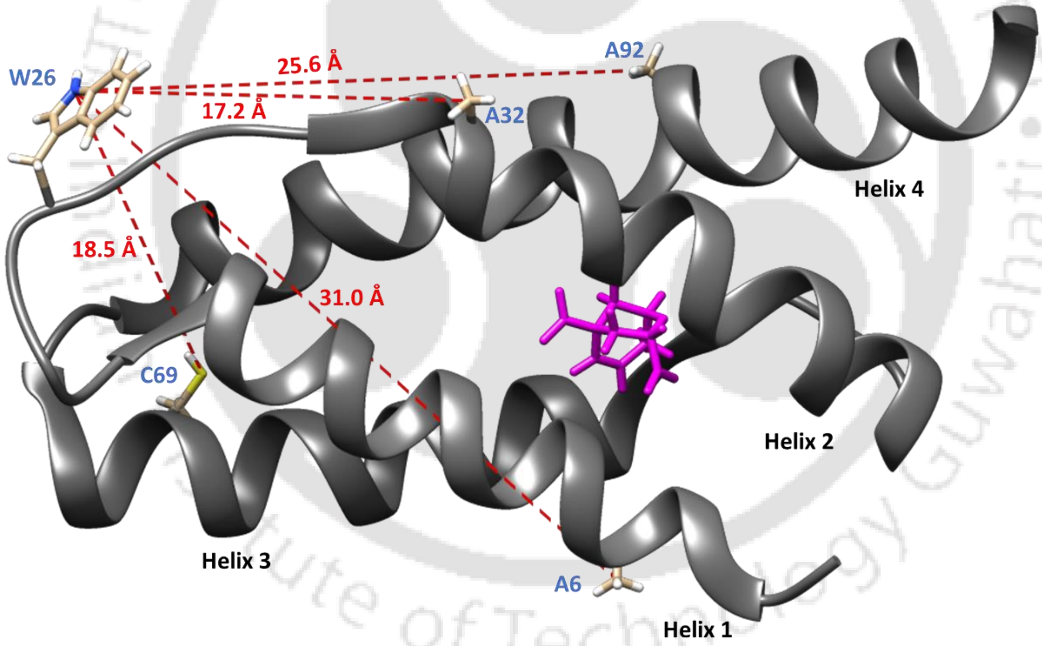
CHAPTER 3

Generation of mutants and purification of mCM

3.1. Selection of sites for mutagenesis

In order to decipher structural and functional information on mCM through fluorescence techniques, the first step required was to generate site-specific mutations at strategic positions on mCM allowing better fluorescence measurements with minimum hindrance to its biological function. The sites for mutagenesis were therefore selected based on a few important considerations. As mentioned earlier in the first chapter, it was important to measure the local transitions in FRET distances between different segments (helices) on mCM in response to ligand (TSA) binding. For this purpose, a single tryptophan residue serving as an intrinsic fluorescence donor and an acceptor probe labeled to a single cysteine residue will be utilized in our studies to acquire sensitive and reproducible information through FRET based experiments (Lakowicz 2006). Details of the FRET technique is provided in Chapter 2. Since mCM contains only two tryptophans (Trp) and a single cysteine (Cys), the above-mentioned approach seemed most suitable to study functional dynamics in mCM. Another important fact to consider is the distance between the donor (Trp) and acceptor (Dansyl) FRET pair, which in turn determines the FRET efficiency between them. For Trp and Dansyl (used in all our FRET experiments), the reported R_0 value for FRET is 22 Å (Lakowicz 2006). Hence, the mCM protein sequence was scanned for sites suitable to be replaced by cysteine in all its helices with due consideration of the distance between tryptophan and the site to be mutated. Through SDM, the single cysteine on mCM will be placed on different helices in different mutants, thus generating multiple single tryptophan and single cysteine variants of mCM. This strategy was devised to facilitate acquisition of data from local/segmental transitions through FRET.

Earlier work by researchers suggest that alanine replacement by cysteine is a well-tolerated mutation in many proteins primarily due to absence of any charged or bulky side chain in alanine (Gregoret and Sauer 1998). Cysteine in proteins also tend to behave slightly hydrophobic unlike the free polar residue (Nagano, Ota, and Y 1999). Hence replacement by the small and hydrophobic alanine residue presents low probability of any drastic structural or functional effect on the protein. Moreover, alanine’s relative abundance, distribution and convenient distance for from Trp for FRET measurements presented alanine as the most suitable candidate for replacement. Therefore, the single cysteine residue (C69, helix 3) in wt mCM was planned to be replaced by alanine in all the mutants, followed by insertion of a single cysteine at the sites A6 (helix 1), A32 (helix 2) and A92 (helix 4) in separate mutants; generating 3 different single cysteine mutants of mCM as shown in **Figure 3.1**. While planning sites for mutation, it was also ascertained that the mutation sites are not in close proximity to the active site of mCM to prevent any possible perturbation in the architecture of the active site or the overall function of the protein. Since mCM contains two closely placed tryptophan residues (Trp 24 and Trp 26), it was also decided to remove one of the tryptophan residues to create a single tryptophan system for less complicated and reliable data interpretation (James et al. 2010) from tryptophan fluorescence experiments (viz. fluorescence lifetime decay, anisotropy, etc). The Trp 24 was replaced by Lys 24 to generate the single tryptophan variant of mCM. This particular replacement was based on previous reports by the Hilvert group (Pervushin et al. 2007) which stated that this mutation didn't alter the catalytic function of mCM and inhibited its tendency to form dimers.



Sequence of wild type mCM disordered enzyme

MIEKL**A^{6}**EIRKKIDEIDNKILKAR**W^{24}**PWAEKLI**A^{32}**ERNS
LAKDVAEIKNQLGIPINDPEREKYIYDRIRKL**C^{69}**KEHN
DENIGIKIFQRLIEHNK**A^{92}**LQKQYLEETLEHHHHHH

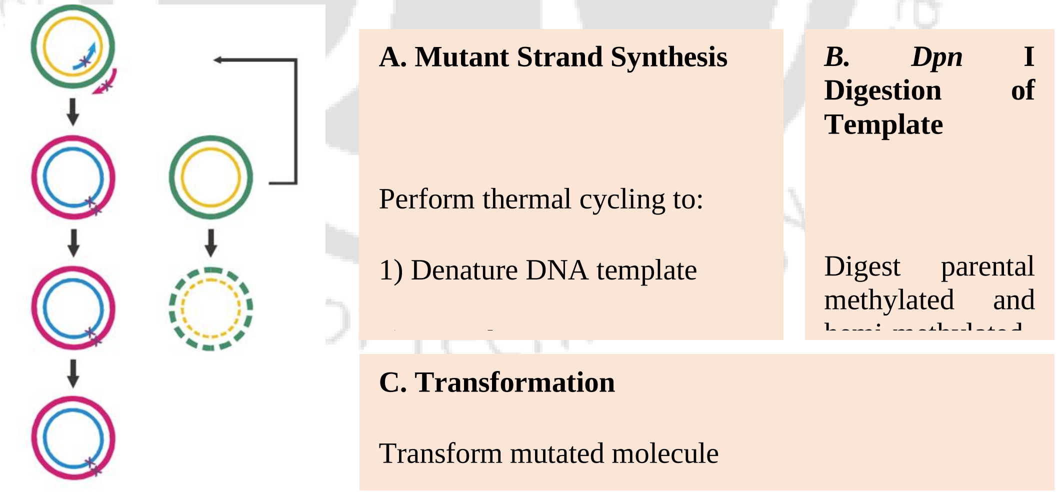
Figure 3.1. Graphical representation of the NMR structure of mCM (PDB: 2gtv) bound to the ligand (TSA) showcasing the sites of mutation and FRET distances. The respective distances from N-atom of indole in Trp 26 to S-atom of Cys 69 and C-atom of all the selected alanine residues (Ala 6, Ala 32 and Ala 92) in the ligand-bound NMR structure are shown in dashed red lines and numbers. The sequence of wild-type mCM is also shown and the sites planned for mutations highlighted in red and numbered. UCSF Chimera 1.13.1 visualization and analysis tool was used for the distance calculations and graphical representation of mCM.

3.2. Generation of mCM mutants

In accordance with the planned sites for amino acid replacements, primers were designed using the Agilent primer design software available online at their website. The primers generated were crossed checked for errors with a standard formula for primer design, as mentioned below:

$$T_{\mathrm{m}} = 81.5 + 0.41(\%\mathrm{GC}) - (675/N) - \%\ \text{mismatch}$$

The primer synthesis was carried out by Integrated DNA Technologies, Inc. The QuikChange II Site-Directed Mutagenesis Kit manufactured by Agilent Technologies was used to generate all the desired mutants. Some minor variations in sample concentrations had to be incorporated for better results. The basic scheme followed for site-directed mutagenesis is shown in **Figure 3.2**



A. Mutant Strand Synthesis

Perform thermal cycling to:

1) Denature DNA template

***B. Dpn* I Digestion of Template**

Digest parental methylated and

C. Transformation

Transform mutated molecule

Figure 3.2. Strategy used for creation of mutation at a single site for amino acid replacement through Site-Directed mutagenesis. Parental plasmid DNA represented as green and yellow concentric circles. Primers and newly synthesized daughter DNA plasmid represented as pink and blue concentric circles. The primers are represented with a arrow head and the site of mutation marked as a small cross in the strands/circles. The digested parent DNA is reprentented by hyped circles.

The detailed method and the conditions used for the PCR-like reaction is mentioned in Chapter 2. Following PCR-like amplification, the mutant plasmids were transformed into DH5α bacterial cells as shown in **Figure 3.3**. The positive colonies obtained in LB agar (Amp^R) plates were further screened for the desired mutations through DNA sequencing (SciGenom Labs). T7 forward and reverse primers were used for sequencing and the resulting DNA sequence (fasta format) was converted to the corresponding amino acid sequence using the ExPASy Translate tool (DNA translation tool). This process was repeated for all the planned mutations.

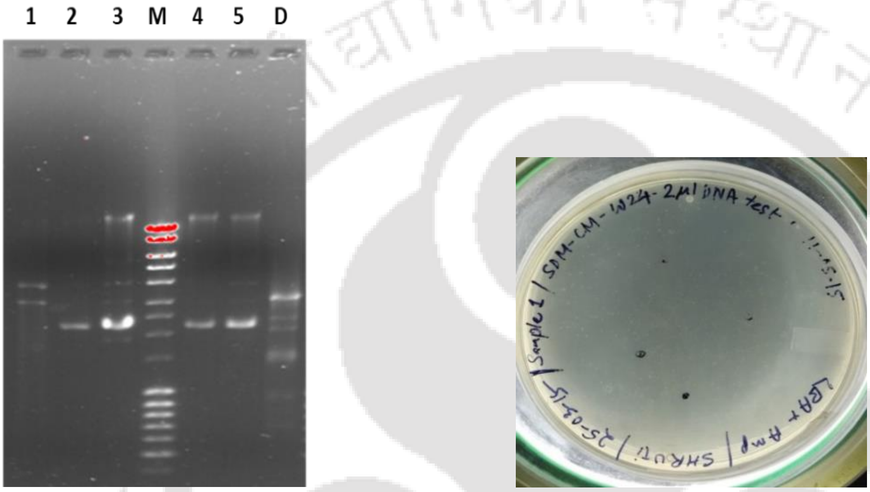


Figure 3.3. *(Left)* Agarose gel showing mutated plasmid mMjCM DNA from positive colonies (1,2,3,4,5) of Mutant 1 (W24K). Parental digested DNA (D) fragments after *Dpn I* digestion can be observed in the gel. *(Right)* Luria-broth agar plates containing ampicillin showing positive colonies of Mutant 1 after PCR-like amplification and transformation into DH5α bacterial cells.

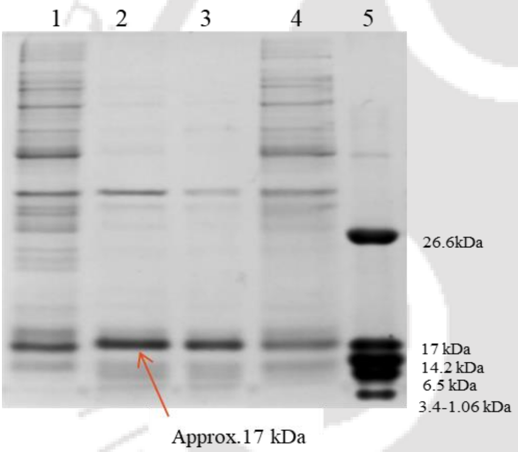
The following **Table 3** summarizes all the successfully generated mutants of mCM with their assigned names that will be used henceforth throughout this thesis.

Table 3. Amino acid residues replaced and their assigned names

mCM	**Mutant1**	**Mutant2**	**Mutant3**	**Mutant4**	**Mutant5**
Amino acid replacement	W24 to K24	W24 to K24 C69 to A69	W24 to K24 C69 to A69 A6 to C6	W24 to K24 C69 to A69 A32 to C32	W24 to K24 C69 to A69 A92 to C92

3.3. Expression of mCM wild-type and the mutant proteins

Wt mCM (wild-type) and the Mutants 1,2 and 3 were expressed in BL21 (DE3) bacterial cells using 0.5 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) for induction of protein expression. Satisfactory levels of protein expression were obtained for Wt mCM as well as all the mutants. The wild-type displayed slightly higher level of protein expression than the mutant variants as shown in **Figure 3.4**. Previous reports (Macbeath, Kast, and Hilvert 1998) show that Wt mCM protein has a molecular mass of approximately 13.2 kDa. Due to the aberrant nature of mobility in SDS-PAGE (Sieber et al. 2011)(Popelka, Uversky, and Klionsky 2014) characteristic of highly charged IDPs, mCM displayed a higher molecular mass corresponding to approximately 17 kDa in 15% SDS-PAGE.



LANES:

1) Uninduced wt mCM sample
2) 0.5mM IPTG induced wt mCM sample
3) 1mM IPTG induced wt mCM sample
4) Induced wt mCM sample pellet
5) Molecular weight marker

Figure 3.4. 15% SDS PAGE gel showing expression of mCM wild-type protein. Some protein can also be observed in the pellet (lane 4).

3.4. Purification of mCM wild-type and the mutant proteins

Wt mCM and all its mutants were successfully purified using the same protocol without any modifications for the mutants as shown in **Figure 3.5**. As mentioned earlier in Chapter 2, Ni-NTA affinity columns (5-10 mL column) were used to purify the His-tagged proteins. All the eluted proteins (from the affinity column) including Wt mCM underwent precipitation during the process of imidazole removal through dialysis. After many trials and errors with variations in the dialysis buffer composition, we found that inclusion of EDTA (100 mM) in the dialysis buffer took care of the problem of protein precipitation. Based on this outcome, we believe that the precipitation was likely caused by the leaching Ni^{2+} ions from the affinity column (Bornhorst and Falke 2000), which may have acted as metal bridges between the proteins. The use of a chelating agent (EDTA) in the dialysis buffer helps to slowly get rid of these metal ions preventing precipitation. The presence of imidazole in the elution buffer helped keeping the his-tagged protein in solution by "masking" the leached Ni^{2+} ions, avoiding the formation of metal bridges. It was crucial to remove imidazole for our experimental studies as it can interference with the tryptophan absorbance and fluorescence measurements due to its high absorbance at 280 nm, overlapping with the absorbance by proteins. A typical yield of approximately 1-1.5 mg/mL was obtained for Wt mCM, Mutant 1 and Mutant 3. The yield of Mutant 2 protein was comparatively less than the other mutants. mCM was found to be "sticky" and bound to commercial protein concentrators. Therefore, the purified proteins were concentrated to a range of 2-3 mg/mL using sucrose beds (for details please refer Chapter 2) for further experiments on mCM. Beyond a concentration of 3 mg/mL, the proteins were precipitated possibly due to aggregation.

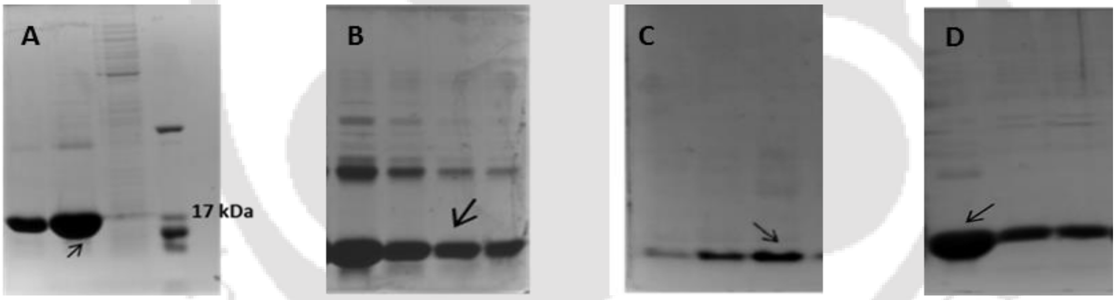


Figure 3.5. 15% SDS PAGE gel showing expression and purification of mCM wild type protein and its mutants (A) wild-type mCM, (B)Mutant 1, (C)Mutant 2, (D) Mutant 3

The Mutants 4 and Mutants 5 were expressed using the same method but further purifications were not carried out. Our initial enzyme activity (Chapter 6) and CD experiments (Chapter 4, Section 4.4) on Mutant1, Mutant2 and Mutant 3 pointed out that the cysteine mutation (C69A) critically affected mCM function (discussed in detail in subsequent chapters). Since Mutant 4 and Mutant 5 both carried this mutation, protein purification and further experiments with these mutants were deemed unnecessary for our studies.

The molecular mass of each of the purified mCM variants were confirmed through mass spectrometry (MALDI). The Wt mCM protein possessed a molecular mass of 13174 Da (**Figure 3.6**) consistent with the previously reported value by Hilvert group (Macbeath, Kast, and Hilvert 1998). The observed molecular mass for the single tryptophan mutant (Mutant1), **Figure 3.7**, the double mutant (Mutant2), **Figure 3.8,** and the triple mutant (Mutant3) **Figure 3.9**, were 13103 Da, 13070 Da and 13101 Da respectively. The expected molecular mass was 13174 Da for Wt mCM, 13116 Da for Mutant1, 13084 for Mutant2 and 13116 Da for Mutant3. A simplified representation of the WT mCM and the three mutants is shown in **Figure 3.10**

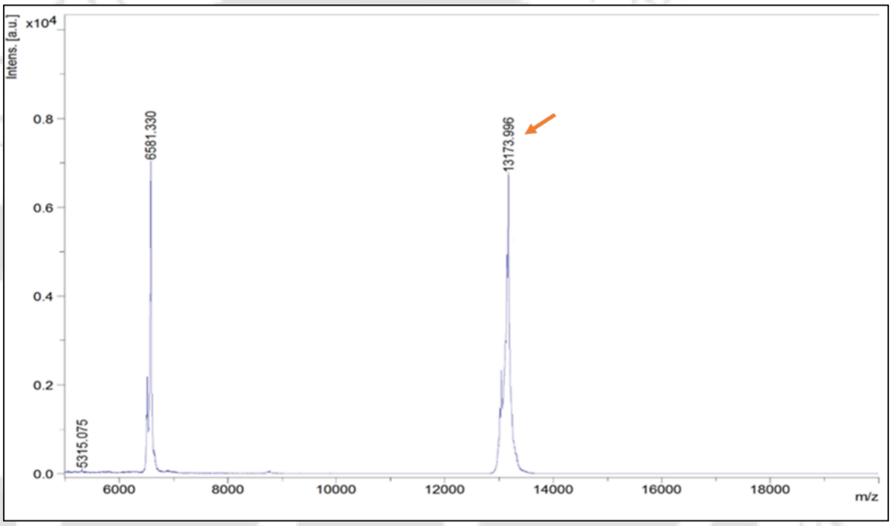


Figure 3.6. The mass spectrum (MALDI) of Wt mCM. Observed mass was 13174 Da.

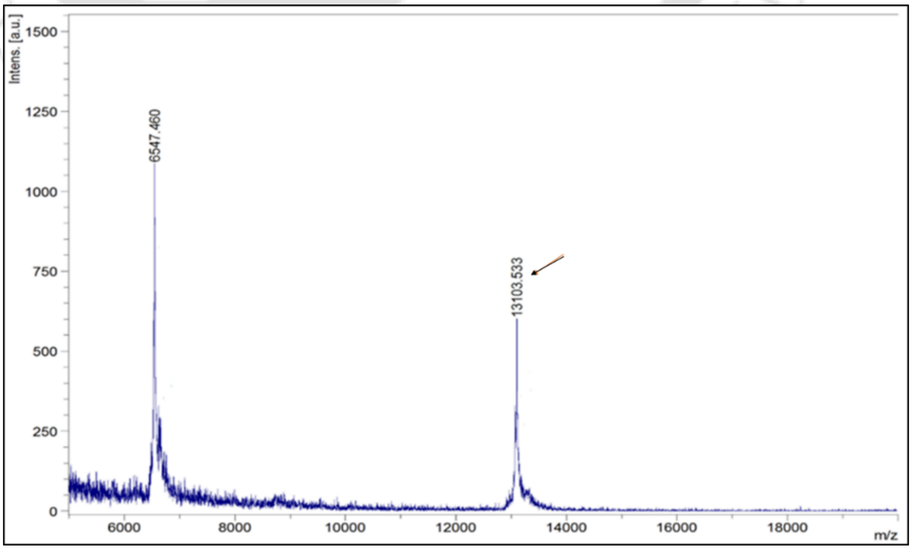


Figure 3.7. The mass spectrum (MALDI) of Mutant1. Observed mass was 13104 Da.

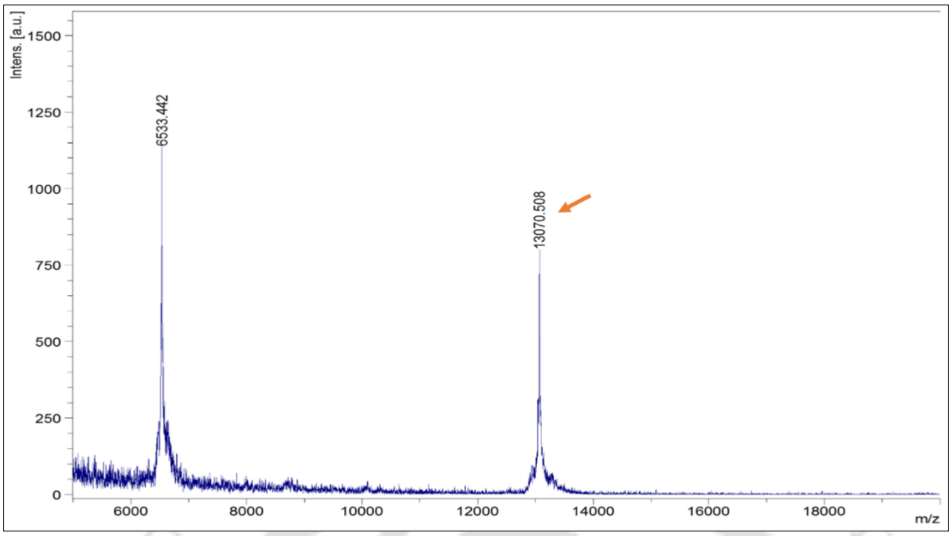


Figure 3.8. The mass spectrum (MALDI) of Mutant2. Observed mass was 13071 Da.

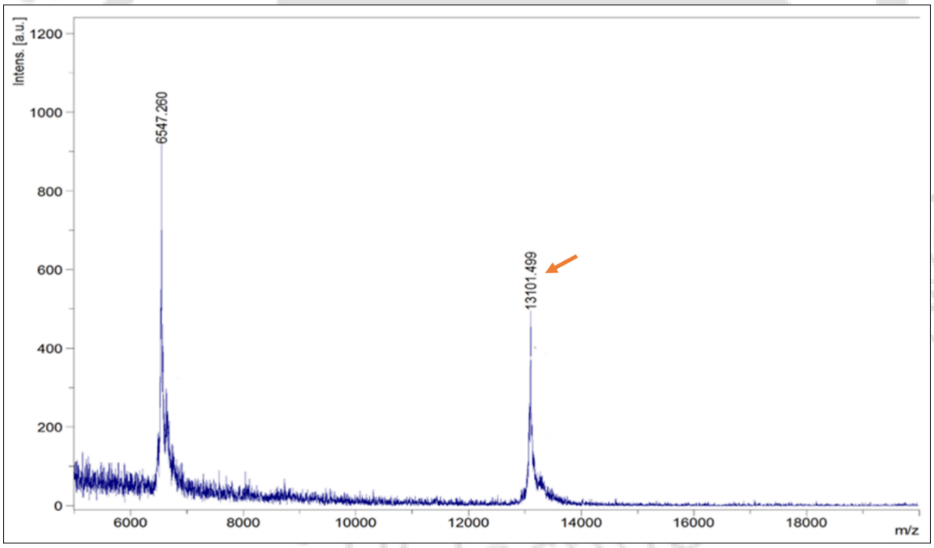


Figure 3.9. The mass spectrum (MALDI) of Mutant3. Observed mass was 13102 Da.

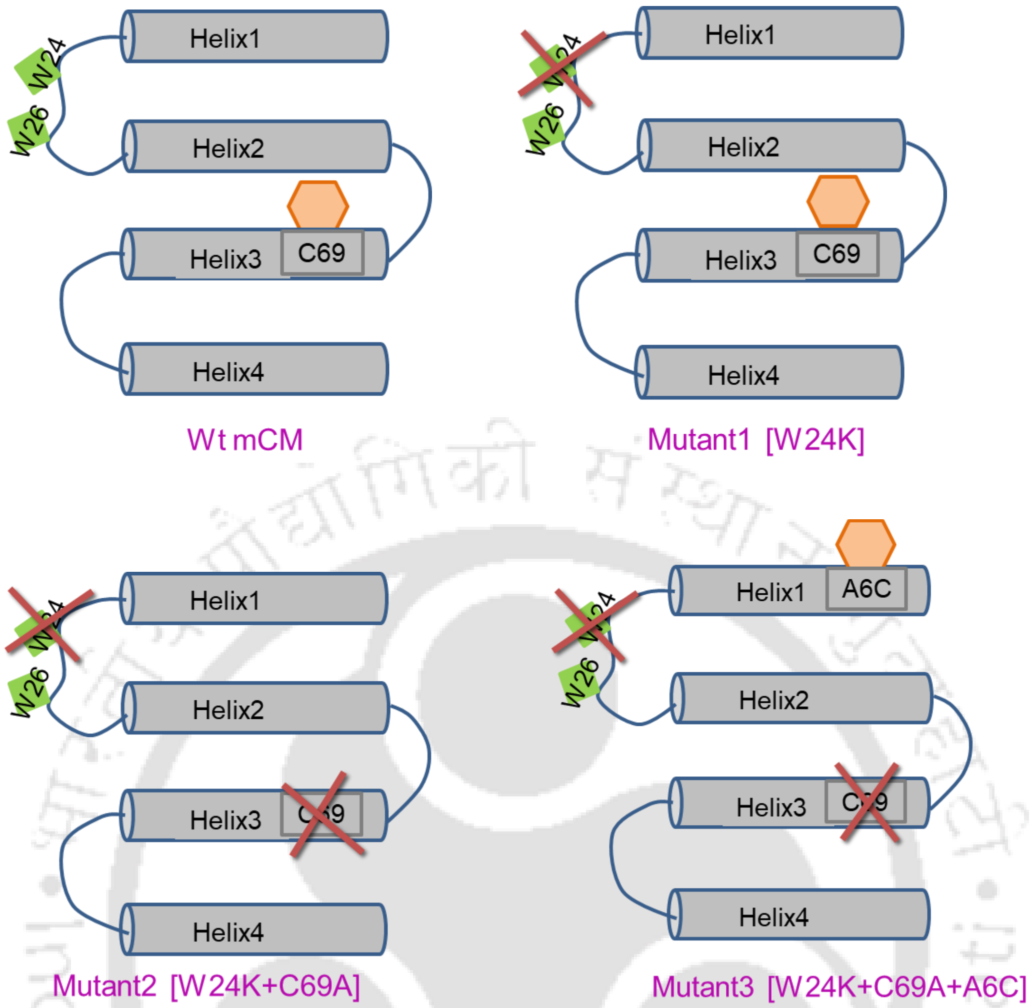


Figure 3.10. Simplified 2-D representation of Wt mCM and its three mutants (Mutant1, Mutant2 and Mutant3) that were expressed and purified for further biochemical and biophysical investigations. The mutated sites in the individual mutants are shown as red crosses. Tryphophan sites are marked as green diamonds. A dansyl probe labeled to the single cysteine for fluorescence studies in mCM is shown as light orange hexagon.

3.5. Conclusions

a) Five mutant variants of mCM with unique Trp-Cys distance pairs were successfully generated using site-directed mutagenesis

b) All the mutants were expressed in a BL-21 (DE3) cells and single bands corresponding to purified Wt mCM, Mutant 1, Mutant 2 and Mutant 3 were observed in SDS-PAGE gels.

c) Wt mCM and its mutants are prone to aggregation at higher concentrations, as observed after purification. Hence, a medium concentration (2-3 mg/mL) was achieved and stored for further experiments.

CHAPTER 4

Steady-state fluorescence and CD measurements of mCM and its mutants

4.1. Intrinsic absorption in wild-type mCM and its mutants

The absorbance spectra of Wt mCM protein displayed the characteristic spectra of tryptophan with a wavelength maximum at 280 nm as shown in **Figure 4.1**. This peak is blue shifted in all the mutants (274-277 nm). This observation arises due to the dominance of Tyr absorbance (three residues in mCM) over the single Trp residue in these mutants.

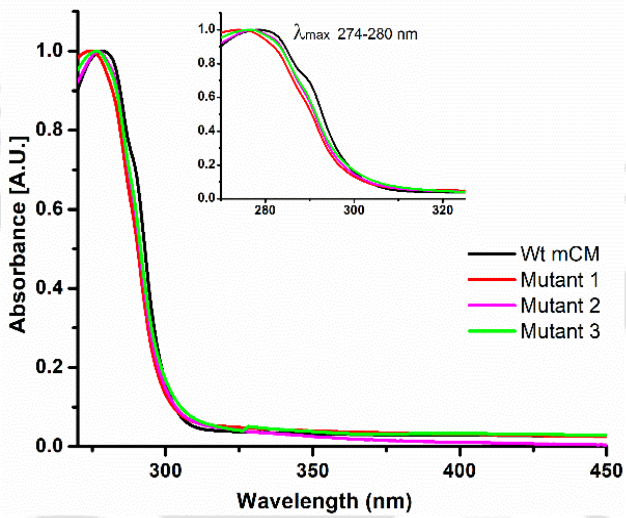


Figure 4.1. Normalized absorbance spectra of mCM and its mutants. [Inset: protein absorbance between 270-325 nm with a wavelength maximum ranging from 274-280 nm.]

4.2. Steady state fluorescence intensity studies of mCM and its mutants

To avoid interference by the Tyr residues, fluorescence spectra of mCM proteins were collected after excitation at 295 nm. A fluorescence peak at 340 nm in Wt mCM suggests that the Trp residues are exposed to aqueous medium as shown in **Figure 4.2**. This is be expected from a protein where the two Trp residues are positioned in a flexible turn connecting two alpha- helices. Absorbance by mCM and its mutants in the 325 nm to 425 nm region is probably due to Protein Charge Transfer Spectra (ProCharTS), which is observed when charged amino/carboxylate groups in the side chains of Lys/Glu act as electronic charge acceptors/donors for photoinduced charge transfer either from/to the polypeptide backbone or to each other (Prasad *et al.*, 2017) (Ansari, Kumar and Ahari, 2018). The single Trp containing mutants displayed a similar profile for the fluorescence emission spectra (λ_{max}341-342 nm). The fluorescence yield for Wt mCM was noticeably reduced compared to its mutant forms and a slight shift of 1-2 nm in the emission maximum was also recorded. The observed difference in fluorescence intensity between the mutants and wild-type might stem from changes in the environment around the single tryptophan (Trp 26) after replacement of the Trp 24 residue in these mutants or possible interactions between the two tryptophans in close vicinity (viz. energy transfer) exclusively in the Wt mCM.

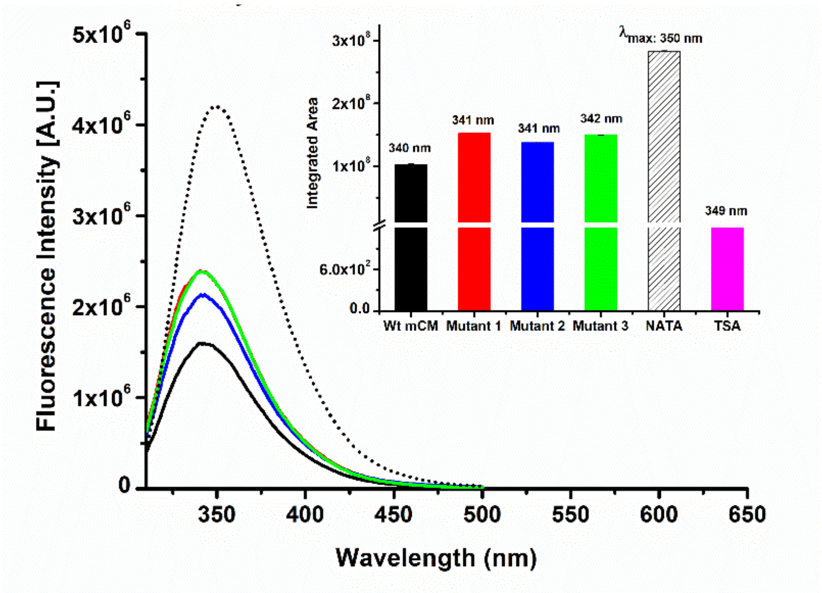


Figure 4.2. Fluorescence spectra (steady-state) of mCM and its mutants obtained after excitation with 295 nm wavelength (excitation slit: 2 nm, emission slit: 5 nm). 20 µM of each of the proteins were used for the experiment. Fluorescence emission of NATA (10 µM) is shown as dotted curve.

In order to ascertain that the tryptophan fluorescence is not directly affected by the ligand, up to 20-fold excess of the TSA (200 µM) w.r.t the Wt mCM (10 µM) was allowed to bind to the protein and the fluorescence was recorded. Addition of TSA had negligible effect on the tryptophan fluorescence intensity in Wt mCM (**Figure 4.3**). This indicates that any structural rearrangement in mCM after TSA binding doesn't alter the flexible turn region or the immediate environment in which the tryptophan(s) resides. A better understanding of the segmental motions and helical compaction after ligand binding was expected from FRET experiments (discussed later). It was also observed that 1mM TSA alone produced fluorescence signals of the order of 10^3 when excited at 295 nm (**Figure 4.3**). Hence, all further fluorescence experiments with TSA was done with concentrations much lower than 1mM to limit the background fluorescence from the TSA.

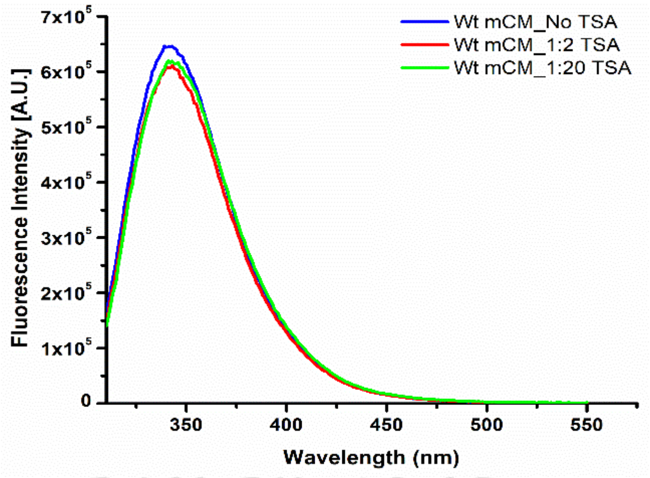
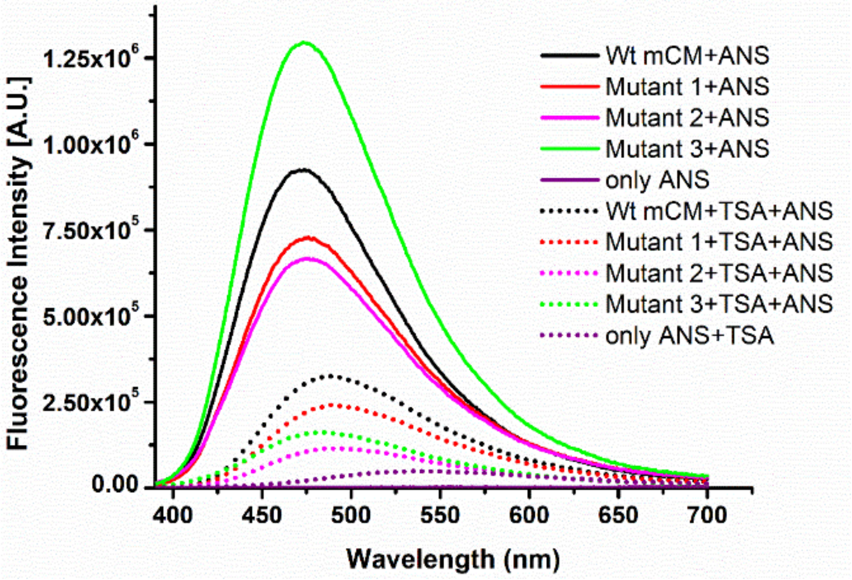


Figure 4.3. Fluorescence spectra of 10 µM Wt mCM (unlabeled with any probe) in presence of 20 µM (1:2) and 200 µM (1:20) TSA.

4.2.1 ANS binding assay of mCM proteins

The TSA binds to the hydrophobic core of mCM ; hence it was quite intended to extract some useful information on the hydrophobic nature of mCM and its mutants in both its free and ligand bound states with the aid of ANS assay (Stryer, 1965). ANS binding to mCM was manifested as a drastic increase in the fluorescence intensity of ANS as shown in **Figure 4.4** . ANS is known to bind preferably to molten globular states of proteins (Ptitsyn *et al.*, 1990) and is therefore expected to be a sensitive tool for our studies.



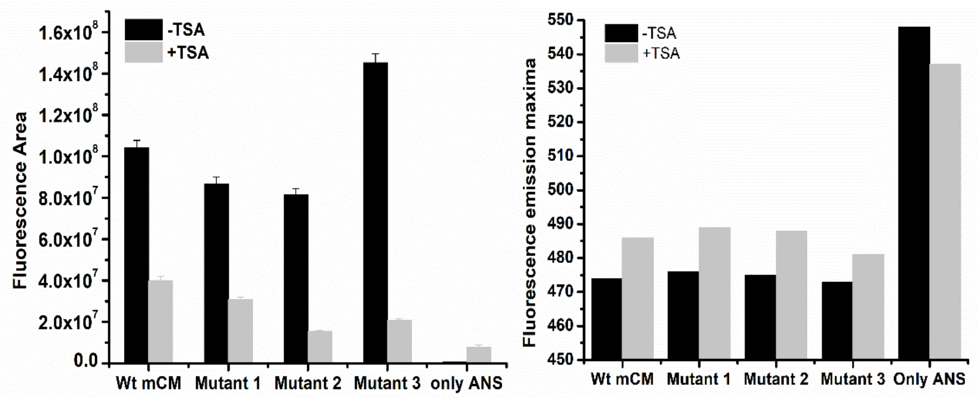


Figure 4.4. (*Top*) The binding of ANS (40µM) to Wt mCM and all the mutants (20µM). Addition of TSA (40 µM) to mCM (20 µM) before ANS binding affected the binding process negatively. An excitation slit of 2 nm and emission slit of 10 nm was used for the measurements and all samples were excited at 375 nm. (*Bottom, left*) Integrated fluorescence area and (*Bottom, right*) emission λ_{max} wavelengths corresponding to the ANS fluorescence intensities for Wt mCM and its mutants in absence and presence of TSA.

ANS displayed higher fluorescence intensity in Wt mCM compared to Mutant1 and Mutant2. The hydrophobic core/patches of Mutant1 and Mutant 2 could have become more solvent exposed or undergone some structural deviation from the wild 3D configuration. Surprisingly, Mutant3 reported even higher fluorescence intensity for ANS. An altogether different structural arrangement of the helices in this triple mutant is possible, giving rise to new and additional hydrophobic patches where ANS could bind. Furthermore, TSA was allowed to bind with the protein before incubation with ANS. This brought about a significant reduction in ANS fluorescence in wild as well as all the variants of the protein. This can happen if TSA binding has rendered ANS inaccessible to its binding sites in the protein. Hence, the TSA and ANS binding sites seem to overlap in the mCM protein. A similar observation was earlier reported by the Hilvert group in ETH Zurich (Vamvaca *et al.*, 2004). This could also mean that a more rigid and tight structure in presence of TSA which closes all entry points for ANS that existed in absence of TSA. Moreover, this study also points to the fact that TSA binds to all forms of mCM, but this binding may not bring about the structural transitions in all forms of mCM necessary for it to carry out its function.

4.2.2 Steady state (ss) FRET measurements

ssFRET data of Wt mCM was recorded to ascertain experimentally the possibility of energy transfer between the tryptophan residues (Trp24/26) in mCM and the dansyl labeled cysteine residue (Cys 69). A significant decline in the fluorescence intensity of Trp in Wt mCM under native condition, was observed after excitation at 295 nm (**Figure 4.5**). To confirm that intramolecular FRET is responsible for this decrease in fluorescence, Gdn.HCl (2-6 M) was used to gradually unfold the protein at 25°C and observe its effect on the tryptophan fluorescence. As expected, an increase in tryptophan fluorescence intensity corresponding to a decrease in the FRET efficiency (0.42 to 0.13) is observed in Wt mCM with increasing unfolding induced by Gdn.HCl (**Table 4.1**). The distance between indoles (donor) of Trp 24 and Trp 26 and dansyl (acceptor) group linked to Cys was calculated to be 23 Å. Since ssFRET is sensitive to sample concentrations, this value is an estimate. Lifetime fluorescence-based FRET measurements for Wt mCM and the mutants were later carried out (please refer Chapter 5) to get more reliable FRET distances (Trp-Cys) for these proteins.

Table 4.1. Change in the intramolecular FRET efficiency of Wt mCM in presence of GdnHcl

Wt mCM in GdnHcl	FRET efficiency
No GdnHCl added	0.43
2M GdnHCl added	0.33
6M GdnHCl added	0.13

The steady-state FRET efficiency for Wt mCM was calculated using the following formula:

$$E = 1 - (F_{DA}/F_D) \qquad 4.1$$

Where F_{DA} is the fluorescence intensity of tryptophan(donor) in presence of dansyl probe (acceptor) and F_D in absence of the acceptor. The λ^{em} range taken for calculating F_{DA} and F_D was 310-400 nm. For FRET efficiency in presence of Gdn.HCl, 20 µM Wt mCM was incubated with 2-6 M Gdn.HCl overnight at room temperature and pH 7.2 before fluorescence measurements were taken.

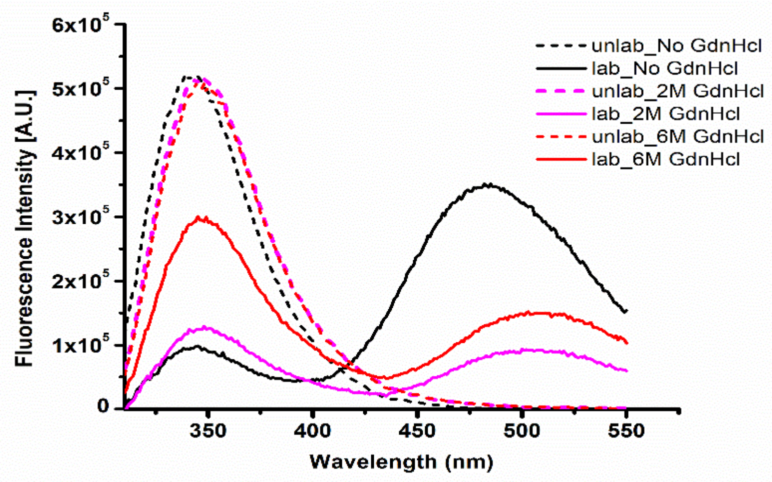


Figure 4.5. Steady state FRET between Trp24/26 and Cys 69 in Wt mCM. Emission spectra corresponding to both Trp and the dansyl probe are observed after excitation at 295 nm (solid lines). Trp in unlabeled mCM is represented in coloured dashed lines. Addition of GdnHcl gradually increases Trp fluorescence in labeled mCM, whereas the fluorescence of unlabeled mCM remains unaffected in presence of Gdn.HCl.

4.3. Steady state fluorescence anisotropy studies of mCM and its mutants

The tryptophan residue(s) on mCM is located on a flexible loop region, connecting two helices. The steady-state anisotropy measurements of Trp in all mCM variants revealed anisotropy values close to 0.1 as shown in **Table 4.2**. Such relatively low values indicate that Trp in mCM can undergo relatively rapid local motions independent of the overall rotation of the protein. Moreover, the mutation of Trp 24 to Lys didn’t affect the rotational freedom or segmental flexibility around the Trp 26 residue. The anisotropy values remained unaltered after addition of the ligand (TSA) to the mCM variants, suggesting that addition of TSA had no influence on tryptophan rotational dynamics.

Table. 4.2. Steady state anisotropy values all mCM variants [λ^{ex} @ 295 nm and λ^{em} @ the respective fluorescence emission maxima for each sample (340-344 nm)]

Sample*	**Anisotropy (r_{ss}) [-TSA]**	**Stdev.**	**Anisotropy (r_{ss}) [+TSA]**	**Stdev.**
Wt mCM	0.111 ± 0.004	0.024	0.106 ± 0.003	0.001
Mutant 1	0.095 ± 0.003	0.003	0.094 ± 0.005	0.002
Mutant 2	0.102 ± 0.004	0.024	0.098 ± 0.003	0.003
Mutant 3	0.109 ± 0.001	0.003	0.096 ± 0.003	0.002
NATA	0.004 ± 0.0005	0.001	0.006 ± 0.0007	0.001

*10 µM sample (each) was taken in PBS buffer pH 7.2 at 25 °C for the measurements.

4.3.1 Trptophan homoFRET in mCM

The two tryptophan residues (Trp 24 and Trp 26) in Wt mCM are located in close proximity to each other, at an average distance of approximately 9 Å as shown in **Figure 4.6**. The R_0 value (the distance corresponding to 50% energy transfer) for Trp-Trp homotransfer is around 6–12Å (Meer et al., 1991). Moreover, the steady-state fluorescence intensity of Wt mCM was visibly lower than all the single tryptophan mutants of mCM (**Figure 4.2**.). This led to investigate the possible existence of Trp to Trp energy transfer in Wt mCM.

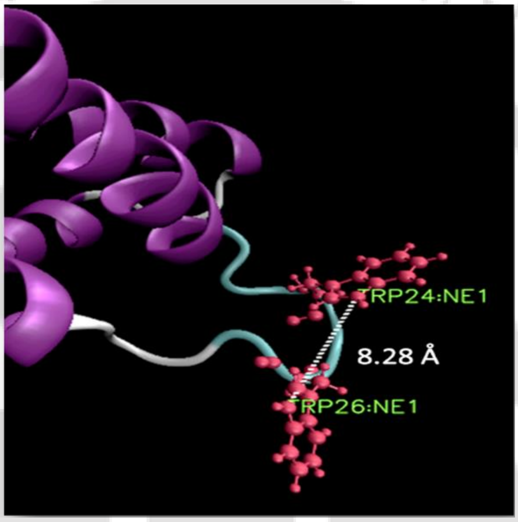


Figure 4.6. NMR structure of Wt mCM focusing on the Trp 24 and Trp 26 residues in a ball and stick representation (red). The distance between the two tryptophan (NE1:NE1) is calculated to be approximately 8.28 Å (VMD program).

In 1960, Weber (Weber, 1960) pointed out that self-transfer (indole to indole or tryptophan to tryptophan) became much less effective upon excitation near the red-edge of the absorption spectrum. This “Weber red-edge effect” has been shown to be a general phenomenon characteristic of homotransfer (Weber and Shinitzky, 1970). For example, in a study of site-directed mutants of fructose 2,6-bisphosphatase, including both single and double tryptophan variants, it was noted that the tryptophan to tryptophan energy transfer could be detected between a specific pair of tryptophan residues by monitoring the polarization of the tryptophan emission upon excitation at two wavelengths, namely 295 nm and 310 nm (Helms *et al.*, 1998). At 295 nm, only tryptophan residues in a protein are effectively excited (i.e., tyrosine excitation is minimized), and the observed depolarization is due to two processes: rotational diffusion of the tryptophan moiety (due to a both “global” rotation of the protein and “local” mobility of the tryptophan) and energy transfer to another tryptophan moiety. Excitation at 310 nm, however, greatly reduces the extent of tryptophan to tryptophan self-transfer due to the Weber red-edge effect so that, at this excitation wavelength, depolarization is primarily due to rotational diffusion processes. Hence, observation of the polarization ratio observed upon excitation at 310 nm and 295 nm (the 310/295 ratio) provides information on tryptophan– tryptophan transfer. In single tryptophan systems, the 310/295 ratios are generally below 2 whereas in multi-tryptophan systems, the 310/295 ratios can be greater than 3 as shown in **Table 4.3**. Therefore, a similar approach was employed to test the presence of homoFRET between the two Trp residues in Wt mCM.

Table 4.3. Anisotropy ratio 310/295 for homoFRET determination

Sample*	**Anisotropy (r_{ss}) @ 295 nm**	**Anisotropy (r_{ss}) @ 310 nm**	**Ratio 310/295**
Wt mCM	0.099 ± 0.004	0.243 ± 0.01	2.42
Mutant 1	0.085 ± 0.005	0.137 ± 0.008	1.61
NATA	0.007 ± 0.001	0.009 ± 0.001	1.12

*10 µM sample (each) was taken in PBS buffer pH 7.2 at 25 ºC for the measurements.

The table above highlights the difference in the 310/295 ratio in the two-tryptophan containing Wt mCM and the single tryptophan variant, Mutant1. A value above 2 (2.42) recorded for Wt mCM suggests the possibility of homoFRET between Trp 24 and Trp 26. Moreover, a fairly blue shifted emission (λ_{max}=340 nm) in Wt mCM favours homoFRET between Trp 24 and Trp 26 (Lakowicz, 2006). As previously mentioned, the relatively lower fluorescence intensity of Wt mCM compared to the mutants also points towards the involvement of some quenching mechanism in Wt mCM. Lifetime fluorescence studies on Wt mCM and the mutants provide us with more proof supporting this observation (refer Chapter 5 and Discussions).

4.4. CD measurements

The visible differences in the fluorescence properties across the mCM variants led us to investigate their secondary structures through Circular Dichroism (CD). A comparison of the CD spectra revealed major alterations in the alpha-helical secondary structure in Mutant 2 and Mutant 3. Mutant 2 exhibited the maximum deviation from the wild type structure. Interestingly, Mutant 3 (the Cys containing triple mutant) possessed more alpha-helicity compared to the cysteine-less double mutant, Mutant 2 (**Figure 4.7**). This highlights the possible role the Cys 69 residue in maintaining the structural integrity of mCM. Analysis of the CD spectra by structural prediction software (K2D3) quantified the decrease in alpha-helicity from 82% in Wt mCM to 54% in Mutant2, which then shows an increase to 59% when the cysteine is replaced in mCM in Mutant3. Mutant 1 was predicted to possess 76% alpha-helicity. We can speculate that local transitions in helix3 of mCM, in which Cys 69 resides, may be necessary for the initial collapse of mCM around its substrate during binding. The Cys to Ala mutation at this site (Cys 69 to Ala 69) might have distorted the helix, handicapping its functional transitions. More insight into this possibility will be addressed in the discussions chapter (Chapter 7).

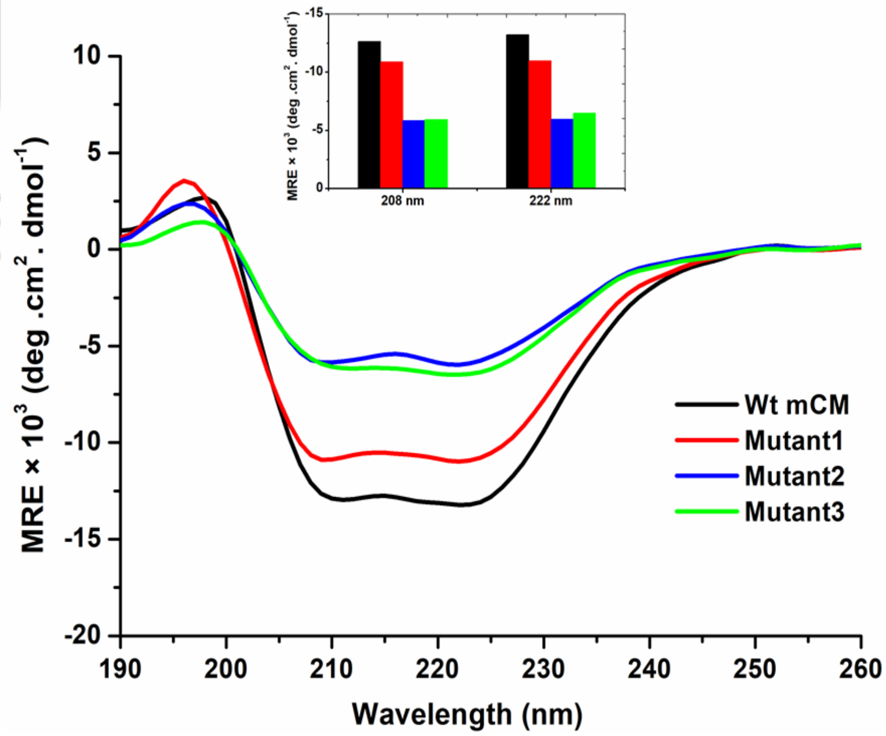


Figure 4.7. CD spectra of mCM and its mutants. A concentration of 10 μM protein in 10 mM phosphate supplemented with 10 mM NaF (pH 7.4) at 25 ^{0}C was used for the CD measurement. (Inset) The values of the two characteristics peaks of α-helices in CD spectra are compared across mCM and its mutants.

The binding of the ligand (TSA) didn’t reveal any major change in the CD spectral profile of mCM or its mutant forms (**Figure 4.8**). These results suggested that TSA binding mainly involved the three-dimensional packing or tightening of mCM around TSA without disturbing or forming (new) secondary structural elements.

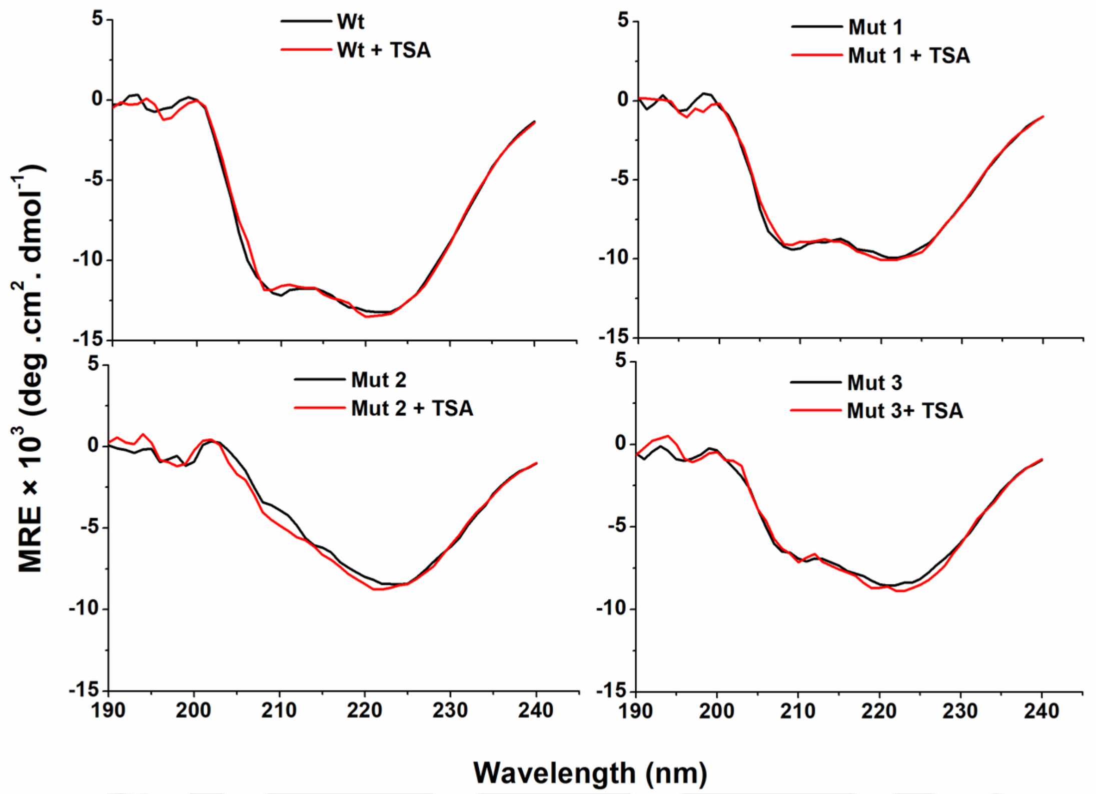


Figure 4.8. CD spectra of mCM variants (15 µM) in presence versus absence of TSA (30 µM).

4.4.1. Chemical denaturation of Wt mCM and its mutants

Chemical denaturation of Wt mCM and its mutants were performed by incubating protein samples in increasing concentration of GdnHCl (0-6 M). Changes in the secondary structure were monitored by following the changes in ellipticity at 222 nm at 25 ^{0}C. The unfolding curve reveals a two-state like transition in Wt mCM as well as the mutants (**Figure 4.9**), but the denatured state is achieved at different GdnHCl concentrations in all the variants of mCM. Wt mCM and Mutant1 seem to have similar stability under GdnHcl denaturation and they both completely lose their secondary structure beyond 4M GdnHCl. This loss in structure is observed earlier at 3M GdnHCl for Mutant3 and still earlier at 2M GdnHCl for Mutant2, the single cysteine mutant. Here, a correlation can be drawn with the secondary structural changes observed earlier in our CD spectra for mCM variants recorded without any denaturants (**Figure 4.7**). Taken together, Mutant2 displayed the greatest loss in alpha helicity having least chemical stability among all mutants.

A sigmoidal fit of the datasets obtained for Wt mCM and the mutants suggest that the mid concentration (C_m) where the protein fraction unfolded is 50% is approximately 2 M for Wt mCM and Mutant1, 1.5 M for Mutant3 and 1 M for Mutant2.

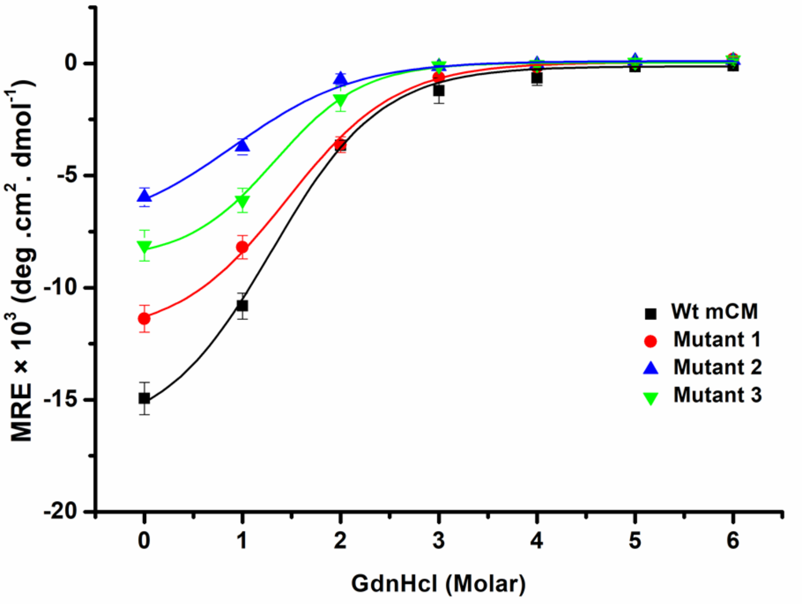


Figure 4.9. Chemical denaturation of Wt mCM and its mutants (10 µM) monitored by UVCD at 222 nm wavelength with increasing concentrations of GdnHCl (0-6 M). Solid fitted lines represent a dose-response curve of non-linear regression analysis using OriginPro 8 software.

An earlier report (MacBeath, Kast et al. 1998) showed that Wt mCM underwent a cooperative transition from folded to unfolded state during urea induced denaturation and the data fits to a two-state transition model, similar to what we observe we GdnHcl induced denaturation.

4.4.2. Thermal denaturation of Wt mCM and its mutants

For thermal denaturation, change in ellipticity of Wt mCM and the mutants were continuously monitored at 222 nm wavelength in the temperature range from 25 to 90 ^{0}C and back from 90 to 25 ^{0}C with a heating/cooling rate of 1 ^{0}C/min. CD spectral scans from 260 nm to 190 nm at specific temperature intervals and their reverse scans were also recorded during the experiment. A similar degree of reversibility after thermal denaturation was observed for Wt mCM and all the mutants (**Figures 4.10 - 4.13**). Interestingly, the thermal denaturation of Wt mCM as well as the mutants appear to be noncooperative in nature and follow a linear transition throughout the temperature scan window. A similar noncooperative thermal denaturation was previously reported for Wt mCM (mMjCM), in absence of any ligand (Vamvaca, Vogeli et al. 2004). They observed modest cooperativity in TSA-bound state whereas the ordered homodimer MjCM enzyme showed a high degree of cooperativity during thermal denaturation. This lack of cooperativity in mMjCM in its free form is argued to be an attribute of its disordered nature.

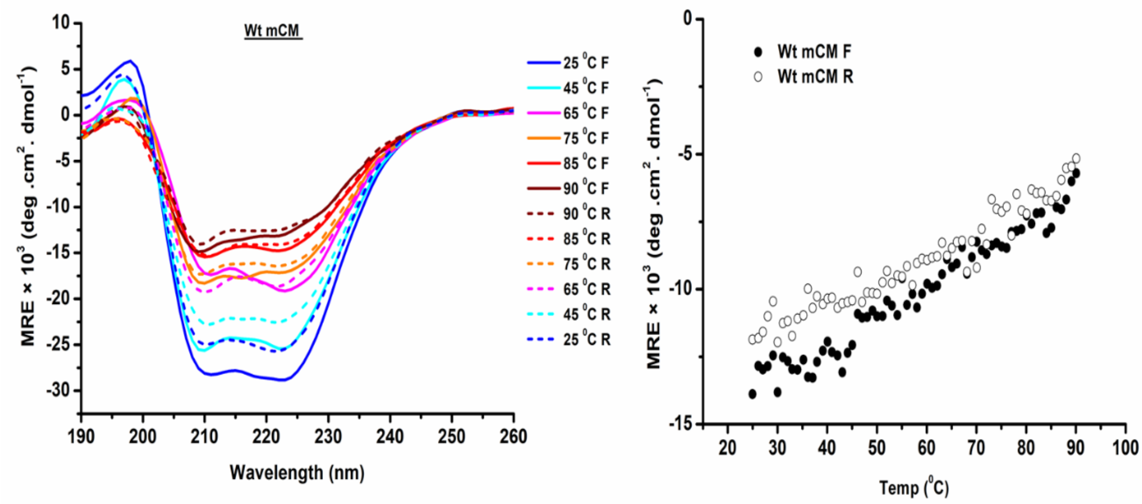


Figure 4.10. Thermal denaturation of Wt mCM. (*Left*) Far UV CD spectra at temperature intervals during the heating cycle (25-90 ^{0}C, solid coloured lines) and the reverse cycle (90-25 ^{0}C, dashed coloured lines) measured in 10 mM phosphate and NaF buffer, pH 7.4, concentration 10 µM. (*Right*) Thermal melting curve of Wt mCM following changes in the ellipticity at 222 nm wavelength. Ellipticity was monitored continuously in the temperature range from 25 to 90 ^{0}C (●) and during a reverse scan from 90 to 25 ^{0}C (○).

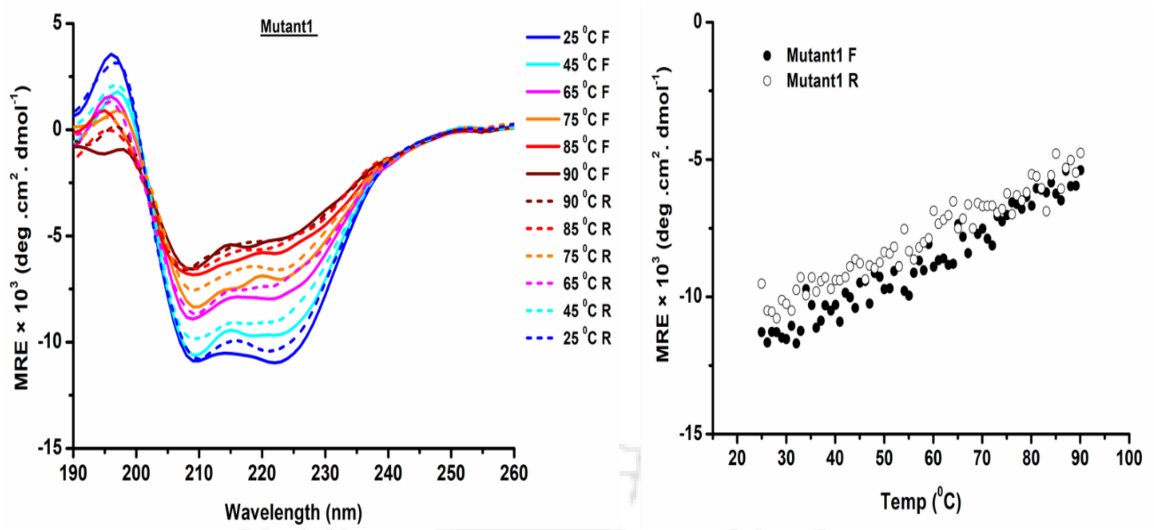


Figure 4.11. Thermal denaturation of Mutant1. (*Left*) Far UV CD spectra at temperature intervals during the heating cycle (25-90 ^{0}C, solid coloured lines) and the reverse cycle (90-25 ^{0}C, dashed coloured lines) measured in 10 mM phosphate and NaF buffer, pH 7.4, concentration 10 µM. (*Right*) Thermal melting curve of Wt mCM following changes in the ellipticity at 222 nm wavelength. Ellipticity was monitored continuously in the temperature range from 25 to 90 ^{0}C (●) and during a reverse scan from 90 to 25 ^{0}C (○).

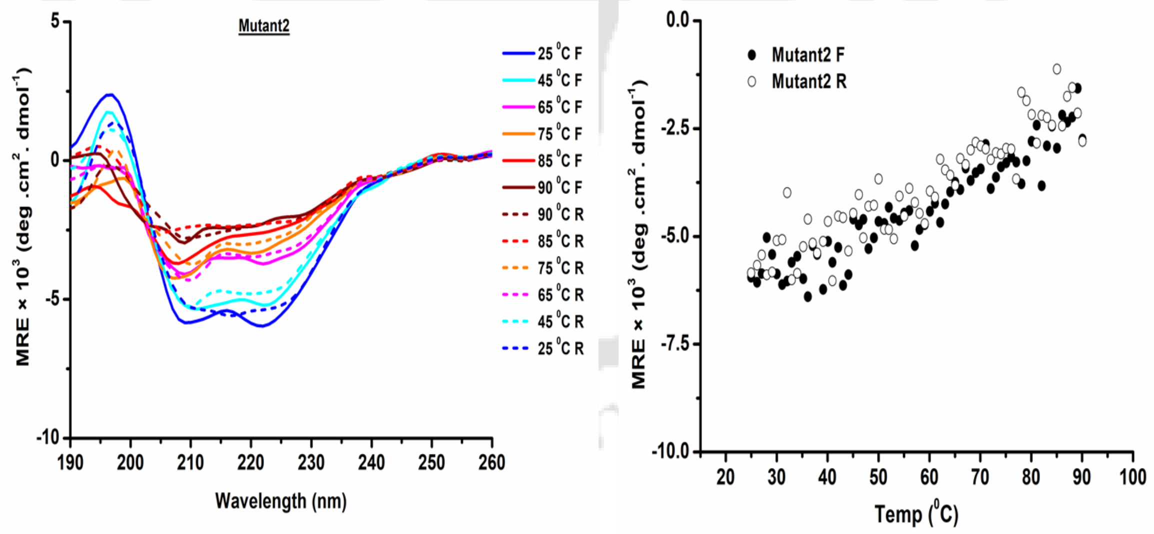


Figure 4.12. Thermal denaturation of Mutant2. (*Left*) Far UV CD spectra at temperature intervals during the heating cycle (25-90 ^{0}C, solid coloured lines) and the reverse cycle (90-25 ^{0}C, dashed coloured lines) measured in 10 mM phosphate and NaF buffer, pH 7.4, concentration 10 µM. (*Right*) Thermal melting curve of Wt mCM following changes in the ellipticity at 222 nm wavelength. Ellipticity was monitored continuously in the temperature range from 25 to 90 ^{0}C (●) and during a reverse scan from 90 to 25 ^{0}C (○).

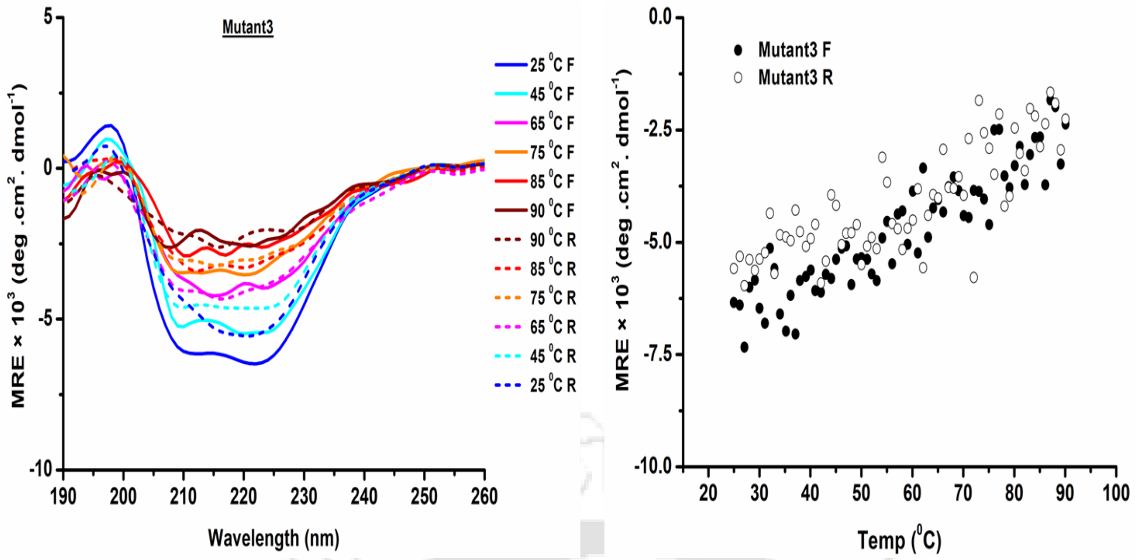


Figure 4.13. Thermal denaturation of Mutant3. (*Left*) Far UV CD spectra at temperature intervals during the heating cycle (25-90 ^{0}C, solid coloured lines) and the reverse cycle (90-25 ^{0}C, dashed coloured lines) measured in 10 mM phosphate and NaF buffer, pH 7.4, concentration 10 µM. (*Right*) Thermal melting curve of Wt mCM following changes in the ellipticity at 222 nm wavelength. Ellipticity was monitored continuously in the temperature range from 25 to 90 ^{0}C (●) and during a reverse scan from 90 to 25 ^{0}C (○).

4.5. Conclusions:

a) The tryptophans in Wt mCM might be involved in homoFRET, affecting their fluorescence intensity and 310/295 anisotropy ratios as compared to its mutants.

b) ssFRET was established between Trp24/26 and dansyl labeled Cys 69, at a FRET distance of 23 Å.

c) Binding of TSA to mCM hindered the binding of ANS to mCM, suggesting overlapping binding sites for both the molecules.

d) The secondary structure of mCM was significantly affected by Cys 69 to Ala 69 mutation in Mutant2 and Mutant3 as revealed by CD spectra.

e) Chemical denaturation of mCM by guanidine hydrochloride reveals varying secondary structural stability of Wt mCM and its mutants; Mutant2 being the least stable. All mCM variants underwent a two-state transition from folded to unfolded state.

f) Thermal denaturation of mCM was noncooperative for all variants and a similar degree of thermal reversibility was observed for Wt mCM as well as the mutants.

CHAPTER-5

Time-resolved fluorescence intensity and anisotropy decay measurements of mCM and its mutants

5.1. Time-resolved fluorescence intensity decay of tryptophan probe in mCM variants

The FRET distance obtained in steady state fluorescence measurements for Wt mCM was verified by time-resolved fluorescence measurements of the tryptophan probe in Wt mCM and the mutant forms. The change in tryptophan lifetime as well as FRET distances in free and ligand bound states of the all variants of mCM were collected and subsequently compared with each other. For ligand bound fluorescence studies with the TSA, concentration of TSA to be used for effective binding studies was calculated from the fractional saturation (Y), defined as the fraction of protein molecules that are saturated with the ligand (TSA). It can be expressed by the following equation:

$$Y = \frac{[L]}{[L]+K_d} \qquad 5.1$$

Where K_d is the dissociation constant for the ligand and [L] is the ligand concentration.

For a ligand concentration of 50 µM (equivalent to the concentration of mCM used for all the lifetime binding experiments) and a dissociation constant (K_d) of 7 µM as reported earlier (Pervushin *et al.*, 2007), the fractional saturation (Y) was calculated to be 0.8. This signifies that approximately 80% of the protein, mCM, is bound to the ligand. Based on this, varying ratios of mCM to TSA ranging from 1:1 to 1:20 were used in all our fluorescence lifetime and anisotropy binding experiments.

5.1.1 Fluorescence lifetime of tryptophans in Wt mCM

Wt mCM featuring two closely placed tryptophan residues reported a mean fluorescence lifetime of 2 ns. Analysis of the fluorescence intensity decay of the tryptophans revealed the presence of three lifetimes (**Table 5.1**). A short component of ~0.8 ns ($\alpha_1 = 0.44$), an intermediate component of ~2.2 ns ($\alpha_2 = 0.40$) and a long component of ~5.2 ns ($\alpha_3 = 0.16$). In presence of 6M Gdn.HCl, the mean lifetime of Wt mCM remains unchanged (~1.9 ns) but individual lifetime components vary significantly **Figure 5.1**. The long component becomes shorter (~3.4 ns) but its amplitude increases from 0.16 to 0.36. The α_i values for the short and intermediate component also decrease signifying changes in environment around the tryptophan residues after unfolding. The presence of two closely placed tryptophans does complicate the interpretation of fluorescence lifetime as contributions from individual tryptophan residues cannot be segregated. The mean lifetime of Wt mCM is drastically different from its mutant variants (~3.3-3.6 ns). This remarkable change has attributed to the existence of homoFRET between the two tryptophan residues in Wt mCM. Quenching of fluorescence resulted in the decrease in fluorescence lifetime as well as steady state fluorescence intensity in Wt CM compared to all its single tryptophan mutants. To test whether the TSA can affect the tryptophan lifetime in Wt CM directly by interacting with the two tryptophan residues (not indirectly through structural changes brought about by FRET due to TSA binding to Wt mCM), an excess of the TSA (1mM) was incubated at a concentration higher than that of the experimental conditions, with Wt CM before conducting tryptophan lifetime measurements. The results (**Table 5.1**) suggested that the TSA doesn't directly interact with the tryptophans and affect their lifetimes. The fluorescence lifetime of the control N-acetyl tryptophan amide (NATA) also remained unaffected in presence of TSA (**Table 5.2**). This provided the confirmation that whatever change in lifetime of tryptophan observed during TSA binding fluorescence lifetime studies can be attributed to structural changes after the TSA binding, and not due to interference in the fluorescence process originating from the tryptophan residues.

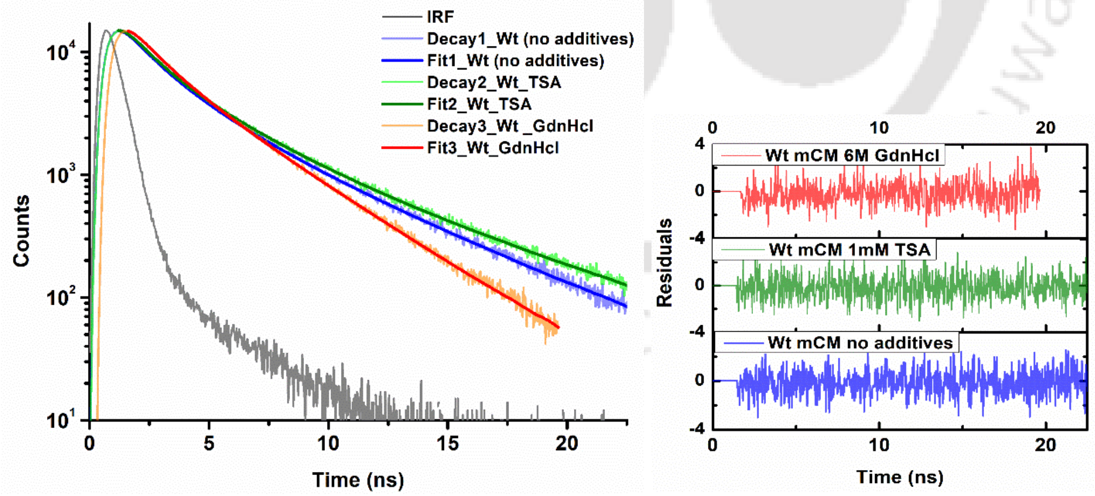


Figure 5.1. Fitted time-resolved fluorescence intensity decay profile of 50 µM Wt mCM with 1mM TSA and 6M GdnHCl compared to the Wt mCM without any additives. Residuals for the fit are shown in the right panels. See Table 5.1 for fitted decay parameters.

The FRET distance between Trp24/26 and Cys69 in Wt mCM was also ascertained by measuring the lifetime of the tryptophan residues in unlabeled versus the cysteine labeled (dansyl probe) form (**Figure 5.2**). The mean lifetime decreased from ~2 ns to ~1.3 ns. The FRET distance was calculated from these values to be ~24Å, close to the value of 23Å obtained in steady state fluorescence experiments. Since steady state fluorescence can be affected by the relative concentrations of the samples, for our future reference, we will consider the FRET distance to be ~24Å.

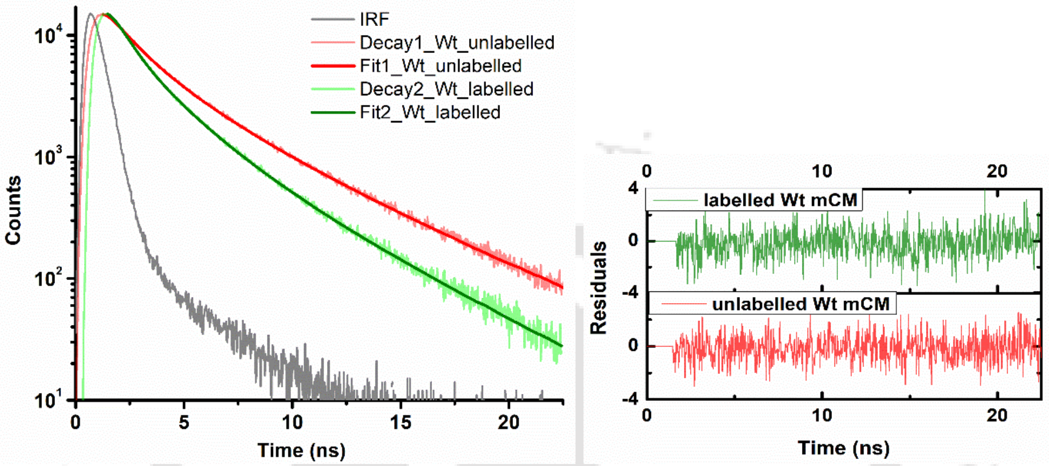


Figure 5.2. Fitted time-resolved fluorescence intensity decay profile of Wt mCM (50 µM) in unlabeled and dansyl (50 µM) labeled forms. Residuals for the fit are shown in the right panels. See Table 5.1 for fitted decay parameters.

Table 5.1. Tryptophan lifetime values (ns) for Wt mCM in PBS buffer pH 7.2 in absence and presence of 6 M Gdn.HCl and 1mM TSA.

Wt mCM	τ_1^{a} (ns)	α_1^{b}	τ_2^{a} (ns)	α_2^{b}	τ_3^{a} (ns)	α_3^{b}	τ_m^{c} (ns)	**Avg.** τ_m^{d} (ns)	χ^{2e}
unlabeled	0.83	0.44	2.2	0.40	5.25	0.16	2.08	2.0 ± 0.2	1.07
unlabeled (+Gdn.HCl)	0.58	0.30	1.48	0.34	3.36	0.36	1.89	1.8 ± 0.1	1.01
unlabeled (+TSA)	0.89	0.49	2.58	0.40	6.5	0.11	2.2	2.05 ± 0.2	1.09
labeled (dansyl)	0.52	0.57	1.68	0.31	4.08	0.12	1.3	1.3± 0.1	1.1

[a]fluorescence lifetime(s); [b]amplitude(s); [c]mean lifetime(s); [d]average mean lifetime(s); [e]reduced χ^2 for the fit.

Table 5.2. NATA lifetime values (ns) in absence and presence of additives

NATA	**Avg. τ^a (ns)**	**χ^{2b}**
No additives	2.8 ± 0.04	1.1
Gdn.HCl (6 M) added	2.7 ± 0.05	1.6
TSA (1 mM) added	2.9± 0.08	1.1

[a]average lifetime; [b]reduced χ^2 for the fit.

The mean lifetimes in Wt mCM after TSA binding didn't reveal any significant change with lower concentrations of TSA (1:1 and 1:2 ratios for Wt mCM and TSA) as shown in **Table 5.3 and Figure 5.3**. A noticeable difference was observed in tryptophan lifetime when an excess of TSA (1:20 ratio) was used. A change to FRET distance from ~24.2 Å to ~22.3 Å was interpreted from the lifetime values (**Table 5.3**). Due to the presence of two tryptophans in proximity with each other involved in energy transfer amongst themselves, a clear picture of the structural transitions in Wt mCM after TSA binding couldn't be derived. So, I proceeded to observe structural transitions in a single tryptophan containing mutant of mCM (Mutant1), as described in the following section (5.1.2).

Table 5.3. Tryptophan lifetime values (ns) for dansyl labeled (at position Cys 69) Wt mCM in PBS buffer pH 7.2 with varying TSA concentrations

Wt mCM	**τ_1^a (ns)**	**α_1^b**	**τ_2^a (ns)**	**α_2^b**	**τ_3^a (ns)**	**α_3^b**	**τ_m^c (ns)**	**Avg. τ_m^d (ns)**	**χ^{2e}**	**FRET distance (Å)**
No TSA	0.52	0.57	1.68	0.31	4.08	0.12	1.3	1.3 ± 0.1	1.1	24.2 ± 0.2
1:1 TSA	0.53	0.62	1.87	0.29	4.59	0.09	1.27	1.3 ± 0.04	1.1	24.1 ± 0.1
1:2 TSA	0.41	0.69	1.68	0.23	4.49	0.07	1.0	1.1 ± 0.09	1.1	22.3 ± 0.2
1:20 TSA	0.48	0.69	1.84	0.24	4.68	0.06	1.0	1.1 ± 0.1	1.1	22.3 ± 0.2

[a]fluorescence lifetime(s); [b]amplitude(s); [c]mean lifetime(s); [d]average mean lifetime(s); [e]reduced χ^2 for the fit.

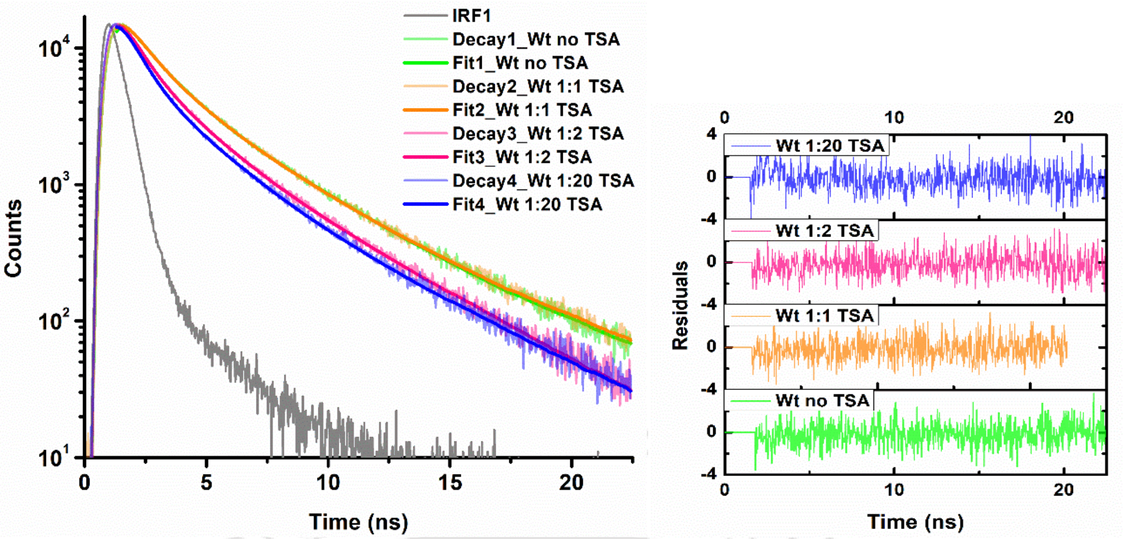


Figure 5.3. Fitted time-resolved fluorescence intensity decay profile of 50 µM dansyl labeled Wt mCM with varying concentrations of the TSA (50 µM, 100 µM and 1mM corresponding to 1:1, 1:2 and 1:20 ratios respectively). Residuals for the fit are shown in the right panels. See Table 5.3 for fitted decay parameters.

The variations of tryptophan lifetime in Wt mCM under different conditions are highlighted in **Figure 5.4**

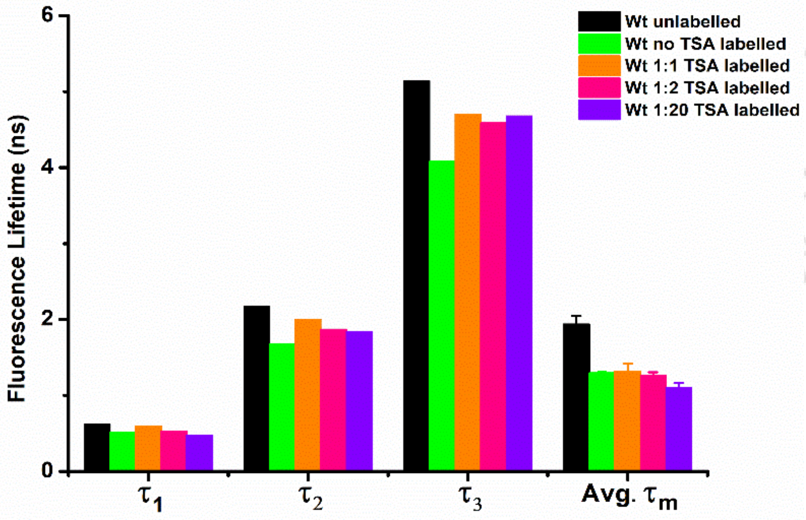


Figure 5.4. Histogram plot highlighting the variations in tryptophan lifetime in Wt mCM in different conditions with or without TSA. The corresponding amplitude for each lifetime value in the histogram in mentioned in Table 5.3.

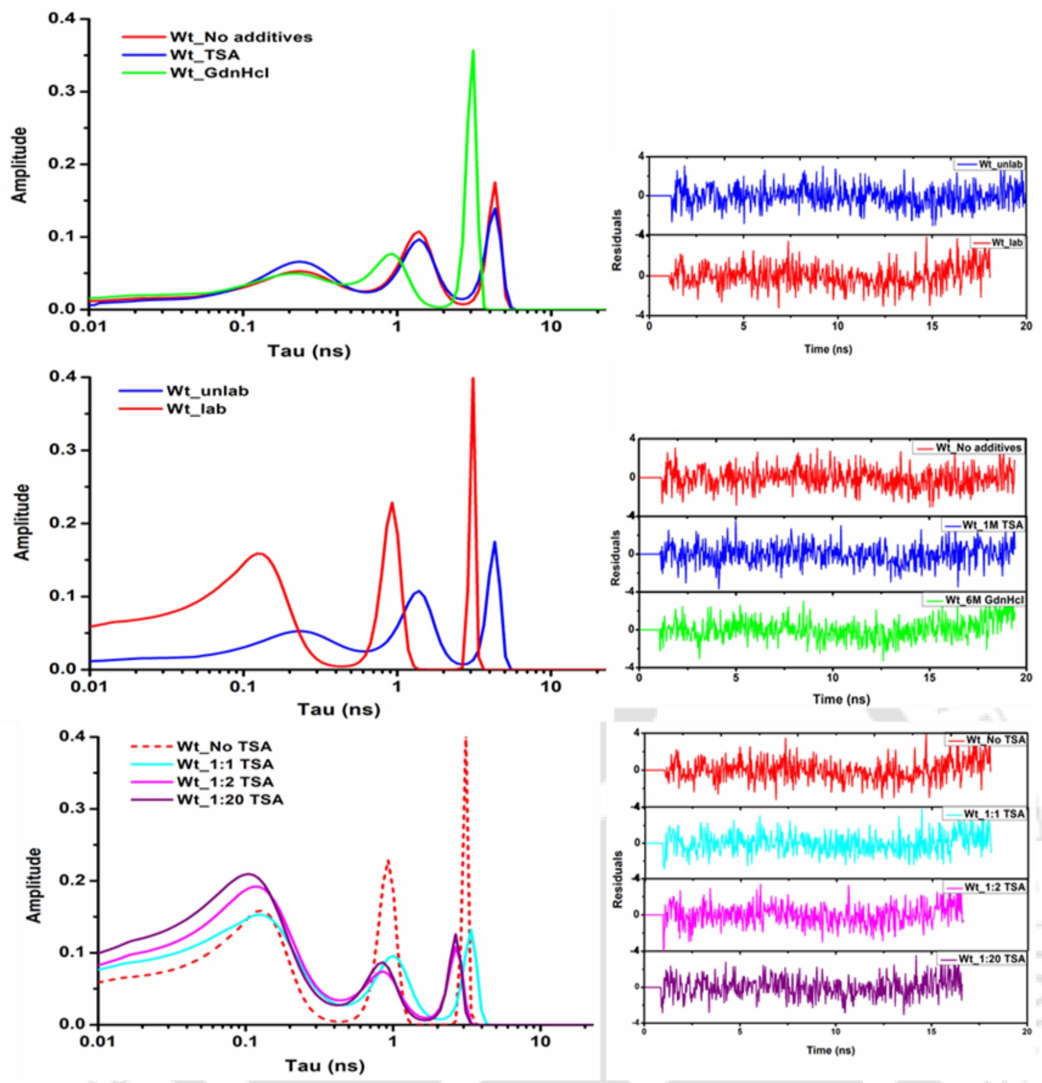


Figure 5.5. MEM distributions for Trptophan fluorescence lifetime in Wt mCM. Top panel: unlabeled protein in absence (no additives) and presence of 1 M TSA and 6 M GdnHCl; Middle panel: unlabeled and dansyl labeled (at Cys 69); Bottom panel: labeled protein in absence and presence of different protein: TSA ratios.

Maximum Entropy Method (MEM) analysis of the Trp fluorescence lifetime data was done for comparison and validation of the discrete analysis method used above. In MEM distributions, the width of a single distribution peak can indicate the uncertainty in that lifetime, implying insufficient information (viz. limited time resolution) in the intensity decay to resolve the distribution any further than already done.

MEM distributions of Wt mCM tryptophan lifetime fluorescence data **(Figure 5.5)**, addition of 1 mM TSA showcased an overlapping distribution of unlabeled Wt mCM without/ with 1 mM TSA. However, the addition of 6M GdnHCl to the protein showed displacement of the lifetime distributions (three components) to lower values, as also observed in the discrete analysis. Dansyl labeled Wt mCM also displayed MEM distribution shift to lower lifetime values. In presence of TSA, the dansyl labeled Wt mCM reported similar lifetime distributions with different Wt mCM: TSA ratios.

MEM distributions mostly overlay on the discrete lifetimes obtained by non-linear least square analysis for Wt mCM as well as the mutants.

5.1.2. Fluorescence lifetime of single tryptophan in Mutant 1

The single tryptophan mutant (W24K), referred to as Mutant 1 in our studies, provided a much simpler and more reliable system to monitor the structural changes in mCM in presence of the TSA. Fluorescence decay analysis of the single tryptophan residue presented two distinct lifetimes and a mean lifetime of ~3.4 ns (**Table 5.4**, **Figure 5.6**). A short lifetime of ~1.8 ns (α_1 = 0.48) and a long lifetime of ~5 ns (α_2 = 0.52) were extracted from the decay fitting and analysis. In presence of 6 M Gdn.HCl, the short lifetime and long lifetime both decreases bringing down the mean lifetime to ~2.5 ns. This decline in the lifetime of tryptophan in 6 M Gdn.HCl may be due to solvent exposure and perhaps quenching by nearby residues as a result of significant structural changes in the vicinity of the single tryptophan moiety after the protein unfolding. The addition of TSA to unlabeled Mutant1 didn't affect the lifetime of tryptophan significantly, as earlier observed in Wt mCM.

Table 5.4. Tryptophan lifetime values (ns) for Mutant1in PBS buffer pH 7.2 under various conditions. 6 M Gdn.HCl and 1mM TSA were used in experiments corresponding to this table.

Mutant 1	$\tau_1{}^a$ **(ns)**	$\alpha_1{}^b$	$\tau_2{}^a$ **(ns)**	$\alpha_2{}^b$	$\tau_m{}^c$ **(ns)**	**Avg.** $\tau_m{}^d$ **(ns)**	χ^{2e}
unlabeled	1.78	0.48	4.97	0.52	3.4	3.4±0.1	1.05
unlabeled (+Gdn.HCl)	1.34	0.53	3.85	0.47	2.51	2.5±0.04	1.06
unlabeled (+TSA)	1.77	0.47	5.1	0.53	3.54	3.5±0.05	0.98
labeled (dansyl)	1.44	0.78	4.41	0.22	2.07	2.0±0.08	1.1

[a]fluorescence lifetime(s); [b]amplitude(s); [c]mean lifetime(s); [d]average mean lifetime(s); [e]reduced χ^2 for the fit.

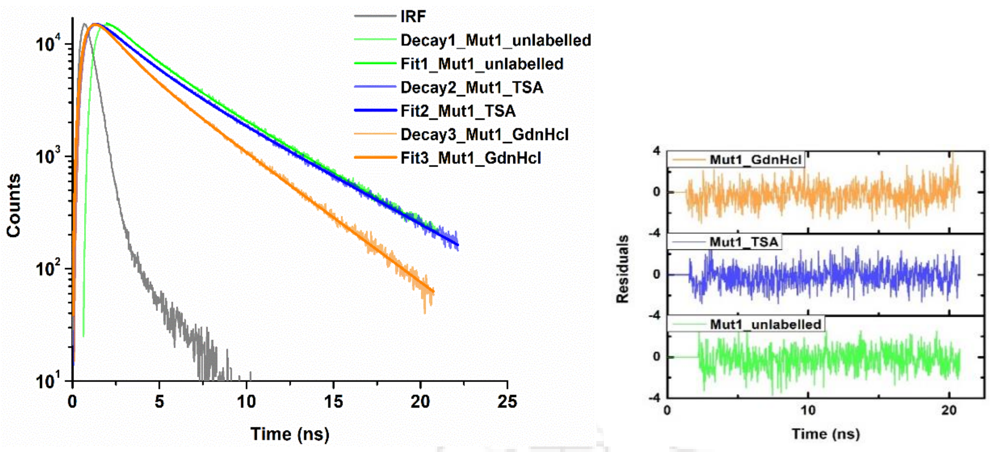


Figure 5.6. Fitted time-resolved fluorescence intensity decay profile of 50 µM Mutant1 with 1 mM TSA and 6 M GdnHCl compared to the Wt mCM without any additives. Residuals for the fit are shown in the right panels. See Table 5.4 for fitted decay parameters.

Labelling (dansyl probe) of Mutant1 at its single cysteine (Cys 69) and monitoring the FRET between the Trp26-Cys 69 through tryptophan lifetime measurement revealed a FRET distance of ~23 Å (**Figure 5.7, Table 5.5**), close to the value (~24 Å) obtained in the two tryptophan Wt mCM. The addition of the TSA to this labeled Mutant1 drastically changed the FRET distance (**Figure 5.8, Table 5.5**). At low TSA concentrations (1:1 ratio and 1:2 ratio with Mutant1), the FRET distance was brought down from ~23 Å to ~18 Å. At slightly higher TSA concentrations (1:4), the FRET distance showed a further decrease to ~16 Å. Addition of even higher concentrations of TSA (1:20) didn't bring any further change to the FRET distance. This drastic change in FRET distance **(Figure 5.9)** can be interpreted as a tightening or collapse of the protein structure around the TSA after its binding, as the flexible loop containing the Trp26 and the helix containing Cys 69 approach each other in the bound form.

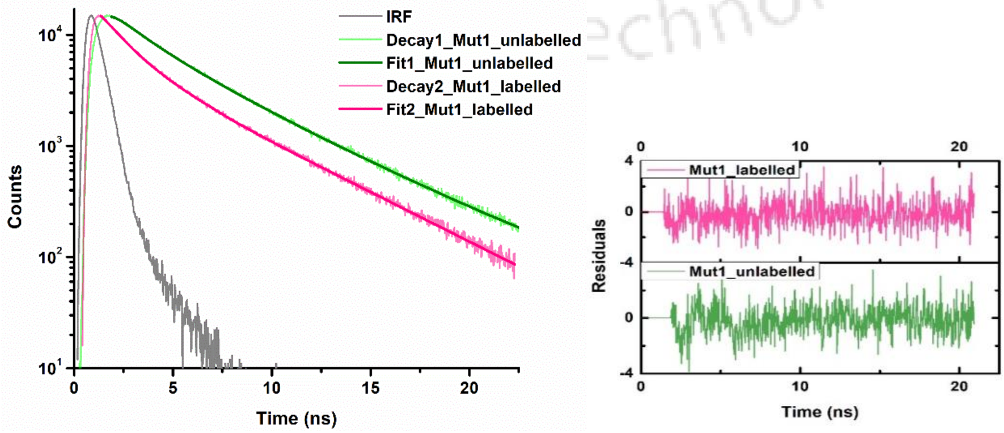


Figure 5.7. Fitted time-resolved fluorescence intensity decay profile of 50 µM Mutant1 in unlabeled and 50 µM dansyl labeled forms. Residuals for the fit are shown in the right panels. See Table 5.4 for fitted decay parameters.

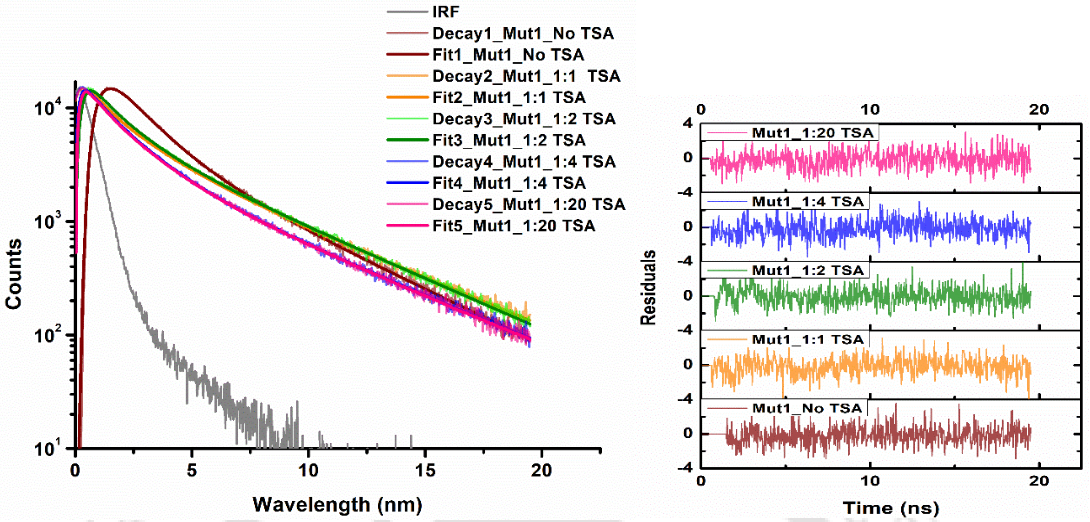


Figure 5.8. Fitted time-resolved fluorescence intensity decay profile of 50 µM dansyl labeled Mutant1 with varying concentrations of the TSA (50 µM, 100 µM, 200 µM and 1mM corresponding to 1:1, 1:2, 1:4 and 1:20 ratios respectively). Residuals for the fit are shown in the right panels. See Table 5.5 for fitted decay parameters.

Another interesting observation is the introduction of a very short lifetime ~0.18 ns in the dansyl labeled form of Mutant1. We assume this lifetime factor to have risen due to FRET and it shows dominance with an amplitude of ~0.7. The amplitude of this very short lifetime component becomes even more dominant (increase in amplitude and further shortening of the lifetime) in the TSA bound forms.

Table 5.5. Tryptophan lifetime values (ns) for dansyl labeled (at position Cys 69) Mutant1 in PBS buffer pH 7.4 with varying TSA concentrations

Mutant1	τ_1[a] (ns)	α_1[b]	τ_2[a] (ns)	α_2[b]	τ_3[a] (ns)	α_3[b]	τ_m[c] (ns)	Avg. τ_m[d] (ns)	χ^2[e]	FRET distance (Å)
No TSA	1.44	0.78	4.41	0.22	---	---	2.07	2.0 ± 0.08	1.1	23 ± 0.2
1:1 TSA	1.75	0.15	5.22	0.09	0.17	0.76	0.85	0.85 ± 0.12	1.1	18 ± 0.3
1:2 TSA	1.5	0.18	4.94	0.11	0.18	0.71	0.93	0.94 ± 0.04	1.1	18 ± 0.2
1:4 TSA	1.62	0.12	5.1	0.04	0.11	0.84	0.49	0.49 ± 0.05	1.1	16 ± 0.5
1:20 TSA	1.61	0.09	5.04	0.04	0.09	0.87	0.42	0.45 ± 0.06	1.1	16 ± 0.3

[a]fluorescence lifetime(s); [b]amplitude(s); [c]mean lifetime(s); [d]average mean lifetime(s); [e]reduced χ^2 for the fit.

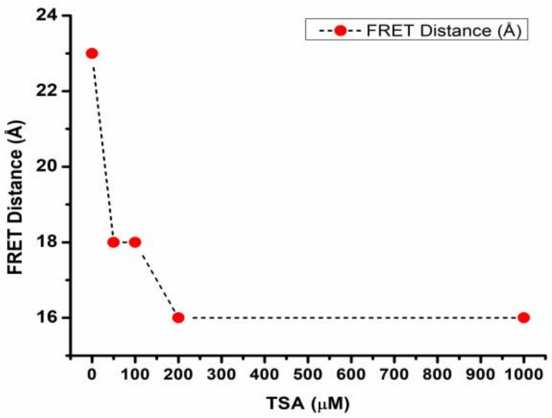


Figure.5.9. FRET distance in Trp26-Cys69 of Mutant 1 observed before and after binding of the ligand TSA (varying concentrations). The exact values with std. dev. are reported in Table 5.5.

The variation of tryptophan lifetime(s) in mutant 1 under different conditions are highlighted in **Figure 5.10**

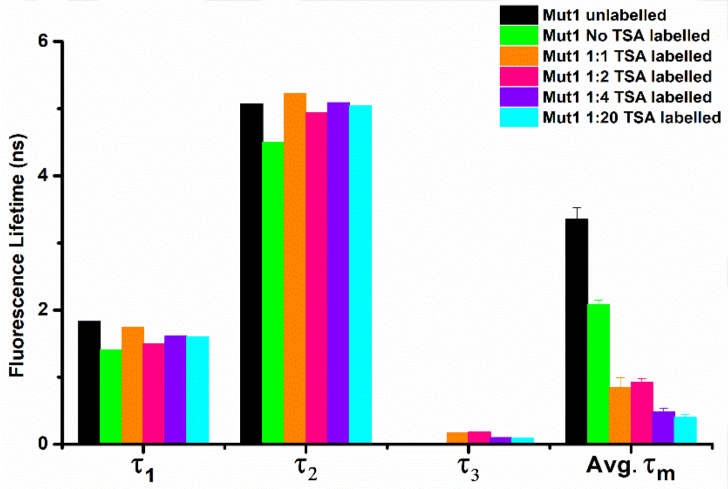


Figure.5.10. Histogram plot highlighting the variations in tryptophan lifetime in Mutant1 in different conditions with or without TSA. The corresponding amplitude for each lifetime value in the histrogram in mentioned in Table 5.5.

MEM analysis (**Figure 5.11**) also revealed significant differences in the lifetime distributions obtained for the higher TSA concentrations (1:4 and 1:20) when compared to TSA unbound Mutant1 protein. The dominant role of a very short lifetime component which increased with higher TSA concentrations was also visible in these distributions, corroborating with the discrete lifetime analysis (~0.18 ns short component, ~0.7 amplitude).

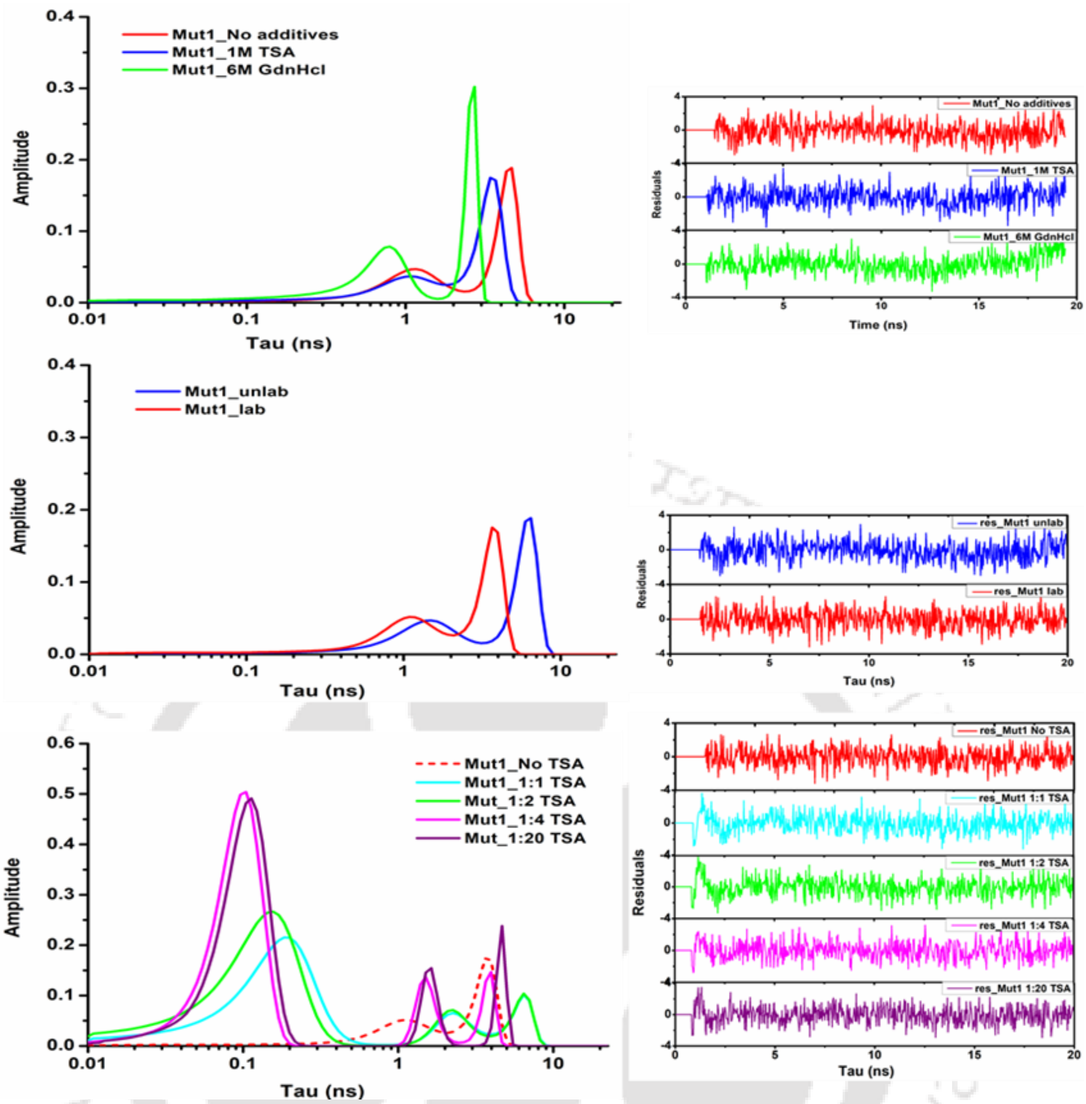


Figure 5.11. MEM distributions for Trptophan fluorescence lifetime in Mutant1. Top panel: unlabeled protein in absence (no additives) and presence of 1 M TSA and 6 M GdnHCl; Middle panel: unlabeled and dansyl labeled (at Cys 69); Bottom panel: dansyl labeled protein in absence/presence of different protein:TSA ratios.

5.1.3. Fluorescence lifetime of single tryptophan in Mutant 2

The double mutant, Mutant2 (W24K/C69A), possesses no single cysteine residue which limits its application for FRET studies by cysteine labelling. Nevertheless, the tryptophan lifetime of Mutant2 was recorded in its unlabeled form and compared to Wt mCM and the mutants. A short lifetime of ~1.7 ns and a long lifetime of ~5 ns were observed with a mean lifetime of ~3.3 ns (**Table 5.6, Figure 5.12**). Interestingly, the two lifetime values obtained for Mutant2 were similar to those of Mutant1 (**Table 5.4**), even though a significant structural and functional alteration was observed through CD and enzyme assay. This may be due to the flexible loop region containing the tryptophan residue being unaffected by the structural changes in Mutant2, also Trp fluorescence senses only the immediate surrounding environment. Lifetime experiments with different concentrations of TSA in unlabeled Mutant 2 reported no noticeable change in tryptophan lifetime (**Figure 5.12** and **Table 5.6**).

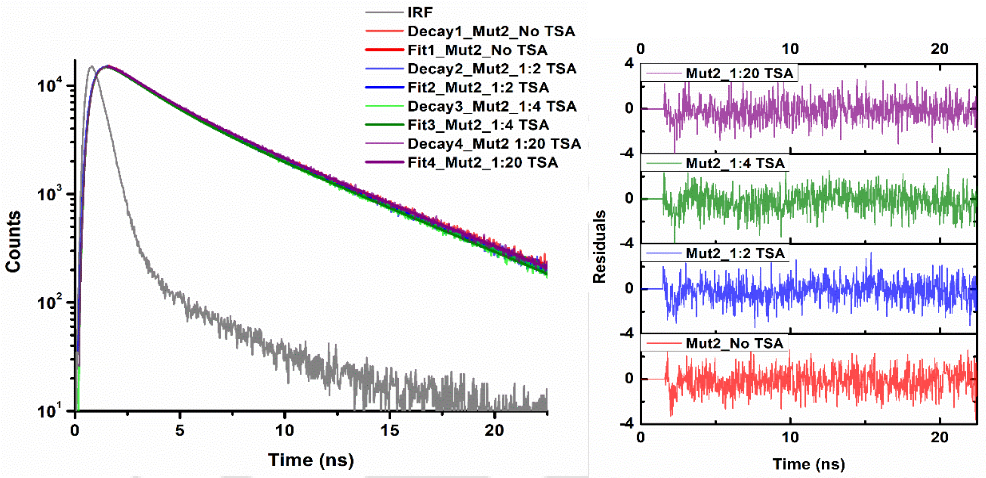


Figure 5.12. Fitted time-resolved fluorescence intensity decay profile of 50 µM unlabeled Mutant2 with varying concentrations of the TSA (50 µM, 100 µM, 200 µM and 1mM corresponding to 1:1, 1:2 and 1:20 ratios respectively). Residuals for the fit are shown in the right panels. Fitted decay parameters are listed in Table 5.6.

Table 5.6. Tryptophan lifetime values (ns) for unlabeled Mutant2 in PBS buffer pH 7.2 with varying TSA concentrations

Mutant2	**τ_1^a (ns)**	**α_1^b**	**τ_2^a (ns)**	**α_2^b**	**τ_m^c(ns)**	**Avg. τ_m^d (ns)**	**χ^{2e}**
No TSA	1.72	0.5	4.97	0.5	3.33	3.35 ± 0.04	1.07
1:2 TSA	1.9	0.56	5.25	0.44	3.38	3.35 ± 0.05	1.1
1:4 TSA	1.87	0.56	5.17	0.44	3.34	3.35 ± 0.06	1.06
1:20 TSA	1.85	0.52	5.16	0.48	3.44	3.40 ± 0.06	1.03

^{a}lifetime(s); ^{b}amplitude(s); ^{c}mean lifetime(s); ^{d}average mean lifetime(s); ^{e}reduced χ^2 for the fit.

In MEM analysis, the lifetime distributions overlapped with each other in all our experimental conditions for the unlabeled Mutant2 (**Figure 5.13**).

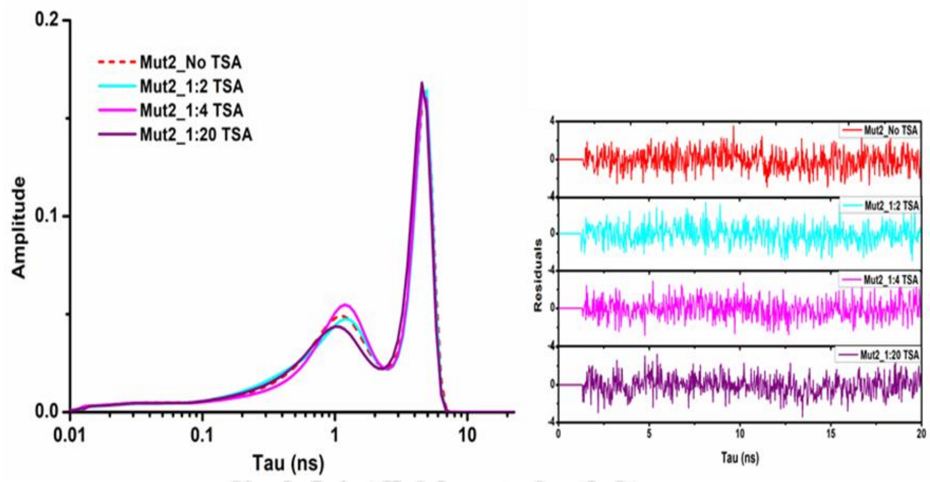


Figure 5.13. MEM distributions for Trptophan fluorescence lifetime in Mutant2. Unlabeled protein in absence/presence of different protein:TSA ratios.

5.1.4. Fluorescence lifetime of single tryptophan in Mutant 3

The triple mutant, Mutant3 (W24K/C69A/A6C) shows a mean tryptophan fluorescence lifetime of ~3.6 ns, a short lifetime of ~2.1 ns (α_1=0.52) and a long lifetime of ~5.2 ns (α_2=0.48), refer **Figure 5.14** and **Table 5.7**. These values are also comparable to the other mutants (Mutant1 and Mutant2) and hence don't point towards any significant change in the environment around the single tryptophan residue. In this case, the FRET distance between Trp26-Cys6 is found to be ~29.3 Å, a distance longer than that reported for Trp26-Cys69. In the TSA bound NMR structure of mCM, the distance between the N-atom of the indole ring in Trp26 and the side-chain C-atom of Ala6 was also reported to be longer (~31 Å) than the Trp26-Cys69 as it is now reported for the free ligand unbound form of mCM.

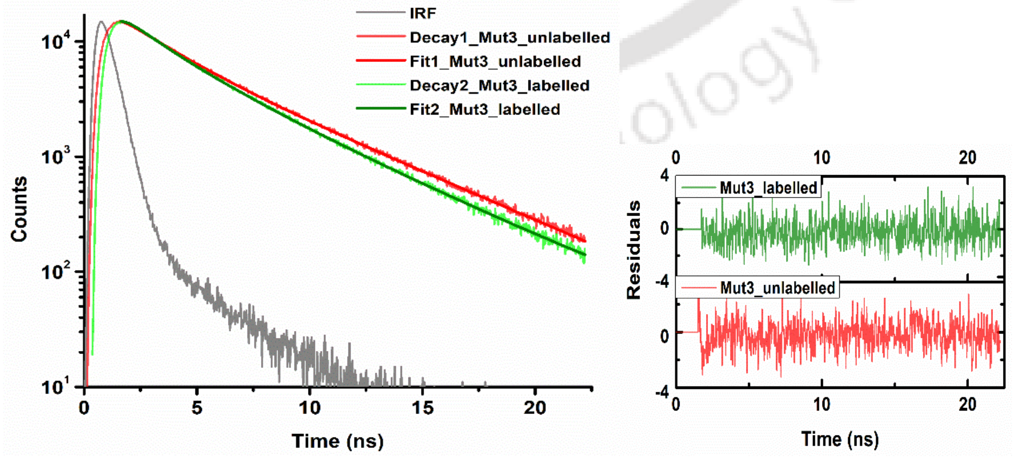


Figure 5.14. Fitted time-resolved fluorescence intensity decay profile of 50 µM Mutant3 in unlabeled and 50 µM dansyl labeled forms. Residuals for the fit are shown in the right panels. Fitted decay parameters are listed in Table 5.7

Table 5.7. Tryptophan lifetime values (ns) for Mutant3 in PBS buffer pH 7.2 in unlabeled and dansyl labeled forms.

Mutant3	τ_1[a] (ns)	α_1[b]	τ_2[a] (ns)	α_2[b]	τ_3[a] (ns)	α_3[b]	τ_m[c] (ns)	**Avg.** τ_m[d] (ns)	χ^2[e]
unlabeled	2.15	0.52	5.2	0.48	---	---	3.6	3.5±0.1	1.03
labeled	2.92	0.49	5.47	0.23	1.1	0.27	3.03	3.0±0.4	1.02

[a]lifetime(s); [b]amplitude(s); [c]mean lifetime(s); [d]average mean lifetime(s); [e]reduced χ^2 for the fit.

In TSA binding experiments (**Figure 5.15**), high concentrations of TSA were required to showcase a noticeable change in FRET distance of Trp26-Cys6 from ~29.3 Å to ~25.7 Å. With increasing concentrations of the TSA, subsequent decrease in the FRET distance was observed. A 20-fold excess of TSA (1:20 ratio), gave an intermediate FRET distance of ~27.2 Å, which decreased ultimately to ~25.7 Å with a 40-fold excess of the ligand (**Table 5.8**). In our enzyme kinetics assay (refer Chapter 6), a similar use of excess enzyme concentration (as compared to the functional Mutant1 and Wt mCM, was required to observe any functional activity in Mutant3. Moreover, CD studies (refer Chapter 4) also indicated significant structural changes in the Mutant3, most probably accounting for weakened TSA binding and major structural changes. Therefore, the FRET distance change observed in this particular mutant may not be a good imitation of the Trp26-Ala6 distance change after ligand binding in Wt mCM.

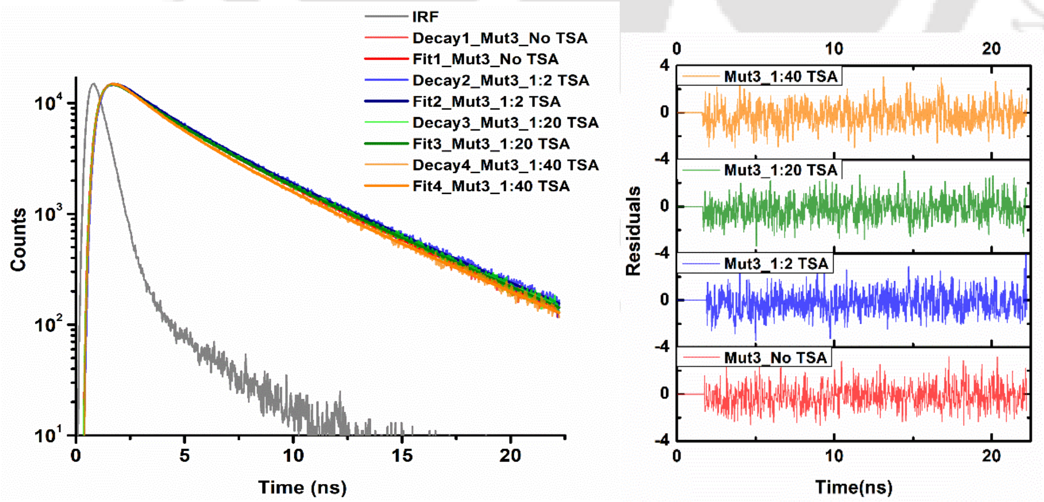


Figure 5.15. Fitted time-resolved fluorescence intensity decay profile of 50 µM labeled Mutant3 with varying concentrations of the TSA (100 µM, 1mM and 2mM corresponding to 1:2, 1:20 and 1:40 ratios respectively). Residuals for the fit are shown in the right panels

The variations of tryptophan lifetime(s) in Mutant3 under different conditions are shown in **Figure 5.16**

Table 5.8. Tryptophan lifetime values (ns) for labeled Mutant3 in PBS buffer pH 7.2 with varying TSA concentrations

Mutant3	τ_1^a (ns)	α_1^b	τ_2^a (ns)	α_2^b	τ_3^a (ns)	α_3^b	τ_m^c (ns)	**Avg.** τ_m^d (ns)	χ^{2e}	**FRET Distance (Å)**
No TSA	2.92	0.49	5.47	0.23	1.1	0.27	3.03	3.0 ± 0.4	1.02	29.3 ± 0.5
1:2 TSA	2.52	0.44	5.23	0.32	0.85	0.23	3.01	3.0 ± 0.06	1.03	29.2 ± 0.6
1:20 TSA	2.47	0.46	5.29	0.29	0.77	0.24	2.87	2.85± 0.02	1.04	27.2 ± 0.4
1:40 TSA	2.23	0.46	5.31	0.25	0.74	0.28	2.6	2.6 ± 0.05	1.03	25.7 ± 0.6

[a]lifetime(s); [b]amplitude(s); [c]mean lifetime(s); [d]average mean lifetime(s); [e]reduced χ^2 for the fit.

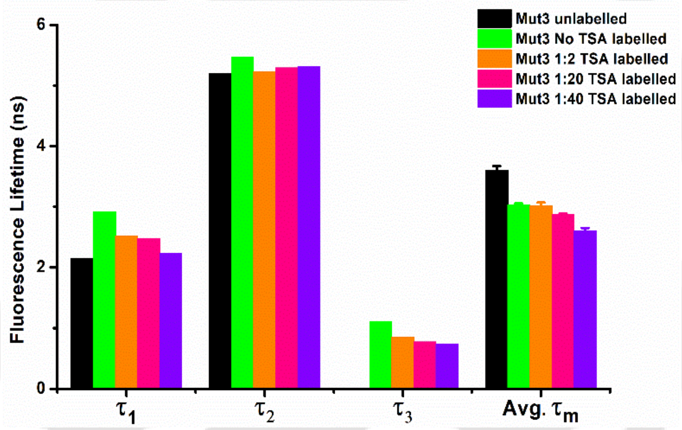


Figure 5.16. Histogram plot highlighting the variations in tryptophan lifetime in Mutant3 in different conditions with or without TSA. The corresponding amplitude for each lifetime value in the histogram in mentioned in Table 5.7 and Table 5.8.

The MEM analysis of the same set of data showed that an excess concentration of the ligand (TSA), 1:40 ratio, presents with visible changes in the lifetime distribution of labeled Mutant3 (**Figure 5.17**). The distribution shows shift to lower fluorescence lifetime values.

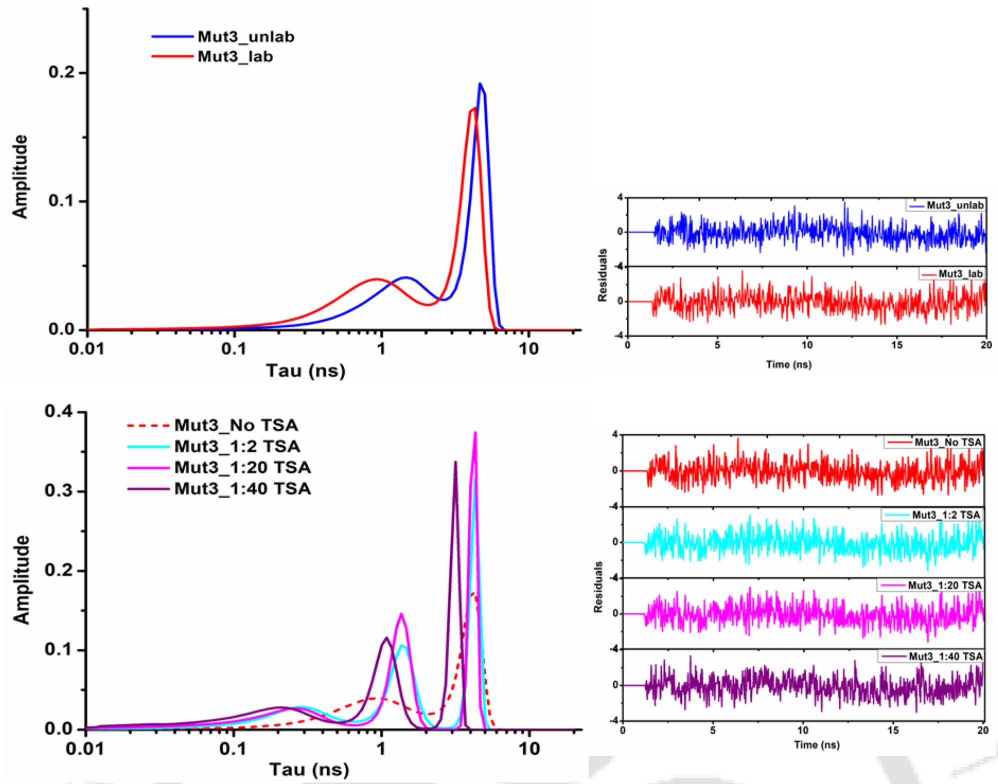


Figure 5.17. MEM distributions for Trptophan lifetime in Mutant3. Top panel: unlabeled and dansyl labeled (at Cys 69); Bottom panel: dansyl labeled protein in absence/presence of different protein:TSA ratios.

A comparison of tryptophan lifetime(s) in unlabeled Wt mCM and its mutants is shown in **Figure 8.18**

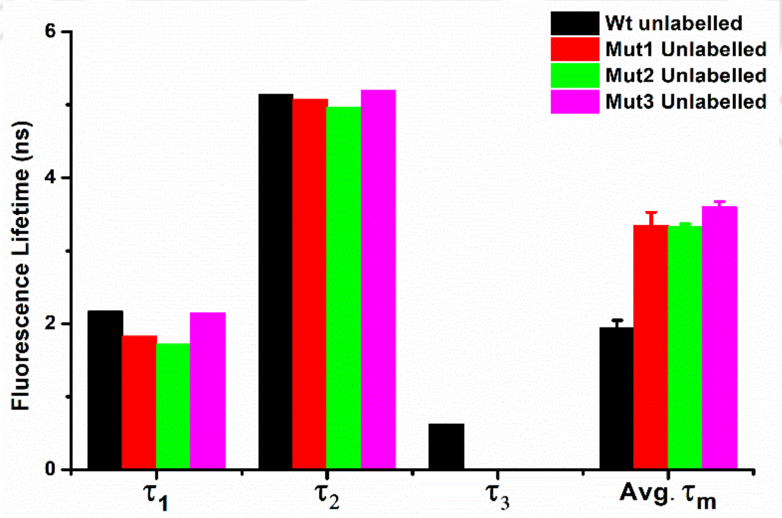


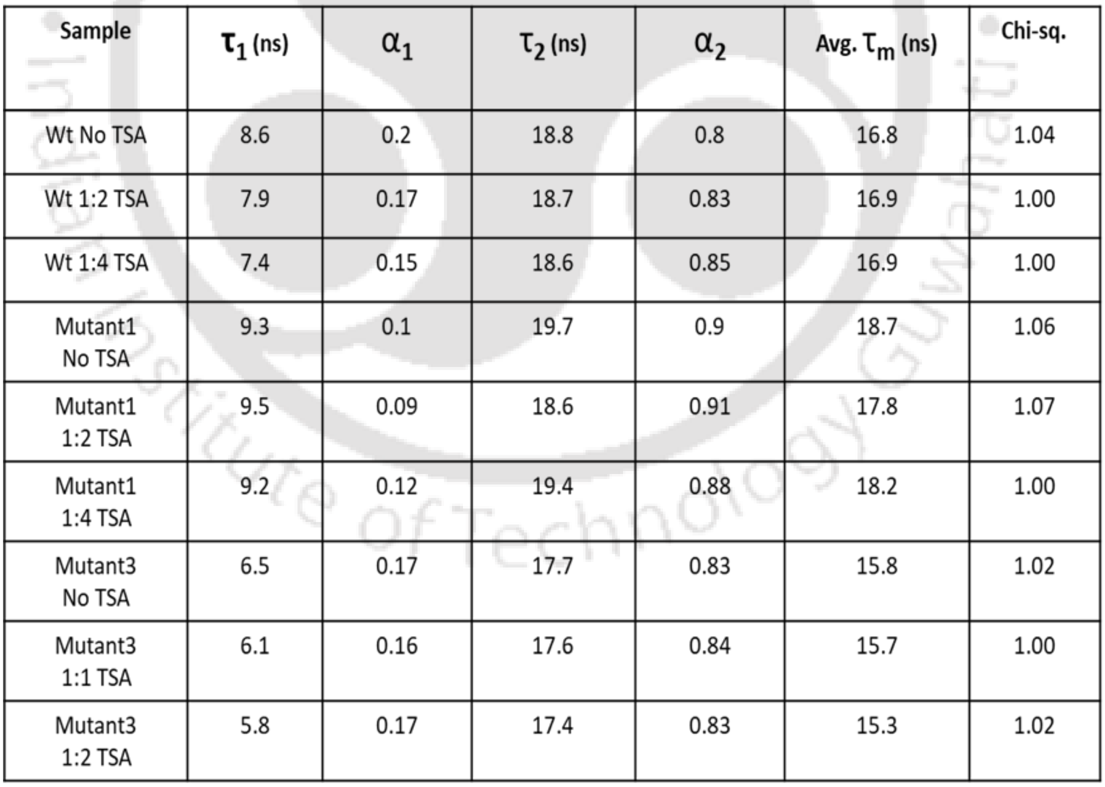
Figure 5.18. Histogram comparing tryptophan lifetime (not fluorescently labeled with dansyl probe, acceptor) across all variants of mCM in different conditions with or without TSA. The corresponding amplitude and std. dev. for each lifetime value in the histogram in mentioned in Table 5.1(Wt mCM), Table 5.4 (Mutant1), Table 5.6 (Mutant2) and Table 5.7 (Mutant3).

5.2. Time-resolved fluorescence decay of cysteine bound dansyl probe in mCM variants

The single cysteine variants of mCM viz. Wt mCM, Mutant1 and Mutant3 were labeled with a dansyl probe, as described earlier in this thesis. The fluorescence lifetime anisotropy decays of the mCM bound dansyl probe under various conditions (with/without ligand) were analysed and then compared to extract valuable information on the conformational dynamics surrounding the probe (indicating the global dynamics of mCM) as well as rotational diffusion rate of the probe. The fluorescence lifetimes of dansyl probe in Wt mCM and its mutants (in absence and presence of TSA) were recorded (**Table 5.9**) to obtain lifetime values for fluorescence decay analysis.

The intrinsic tryptophan in mCM, having a shorter fluorescence lifetime (~ 2 ns), was not used for our anisotropy decay studies exceptt for Mutant2, which lacks any cysteine site for the dansyl probe labeling.

Table 5.9. Fluorescence lifetimes of dansyl probe in Wt mCM and its mutants in absence and presence of the ligand TSA.



Sample	τ_1 (ns)	α_1	τ_2 (ns)	α_2	Avg. τ_m (ns)	Chi-sq.
Wt No TSA	8.6	0.2	18.8	0.8	16.8	1.04
Wt 1:2 TSA	7.9	0.17	18.7	0.83	16.9	1.00
Wt 1:4 TSA	7.4	0.15	18.6	0.85	16.9	1.00
Mutant1 No TSA	9.3	0.1	19.7	0.9	18.7	1.06
Mutant1 1:2 TSA	9.5	0.09	18.6	0.91	17.8	1.07
Mutant1 1:4 TSA	9.2	0.12	19.4	0.88	18.2	1.00
Mutant3 No TSA	6.5	0.17	17.7	0.83	15.8	1.02
Mutant3 1:1 TSA	6.1	0.16	17.6	0.84	15.7	1.00
Mutant3 1:2 TSA	5.8	0.17	17.4	0.83	15.3	1.02

5.2.1 Time-resolved fluorescence anisotropy decay of dansyl probe in Wt mCM

The rotational correlation time (θ) of Wt mCM was calculated by measuring the fluorescence anisotropy decay of its dansyl probe at pH 7.4 in PBS buffer in absence or presence of varying concentrations of the ligand (TSA). **Table 5.10** summarized the change in values of the correlation time(s) obtained after addition of TSA to Wt mCM. The presence of two correlation time is observed after analysis of the anisotropy decay profile (**Figure 5.19**). The unbound Wt mCM (30 µM) displayed a fast rotational component of 1.32 ns with amplitude of 0.44 and a slow rotational component of 8.6 ns with amplitude of 0.56. The fast component corresponds to the rotational flexibility around the dansyl probe, whereas the slow component relates to the global rotation of the entire protein.

Table 5.10. Anisotropy decay data of dansyl labeled Wt mCM (-/+ TSA) in PBS buffer pH 7.4

Wt mCM	**r_0[a]**	**r_{ss}[b]**	**θ_1[c] (ns)**	**β_1[d]**	**θ_2[c] (ns)**	**β_2[d]**	**χ^2[e]**
No TSA	0.18	0.013	1.32 ± 0.04	0.44	8.6 ± 0.15	0.56	1.2
1:2 TSA	0.19	0.012	0.71 ± 0.01	0.5	7.6 ± 0.09	0.5	1.2
1:4 TSA	0.17	0.013	0.44 ± 0.01	0.61	6.35 ± 0.06	0.39	1.3

[a] initial anisotropy; [b] steady state anisotropy calculated from fit; [c] rotational correlation time(s); [d] fractional amplitude associated with the correlation time; [e] reduced χ^2 for the fit.

A sharp decrease in the value of the fast component to 0.71 ns (1:2 ratio, 60 µM TSA) and further down to 0.44 ns (1:4 ratio, 120 µM TSA) suggested that the rotational motion of the dansyl probe in the TSA bound form becomes faster. The increase in the corresponding amplitude (0,44 to 0.61) for the shorter rotational time in TSA bound Wt mCM also implies increased freedom for rotational motion for the dansyl probe, after binding. The slow component decreased in TSA bound form (**Table 5.10**) indicating faster global rotation due to compact structure of Wt mCM after ligand binding. The surprising change in flexibility of the dansyl probe in bound state could indicate an active role of the amino acids surrounding the dansyl probe (position Cys 69) in ligand binding. A comparison of the correlation time(s) in dansyl labeled Wt mCM in its unbound and TSA bound states are shown in **Figure 5.20.**

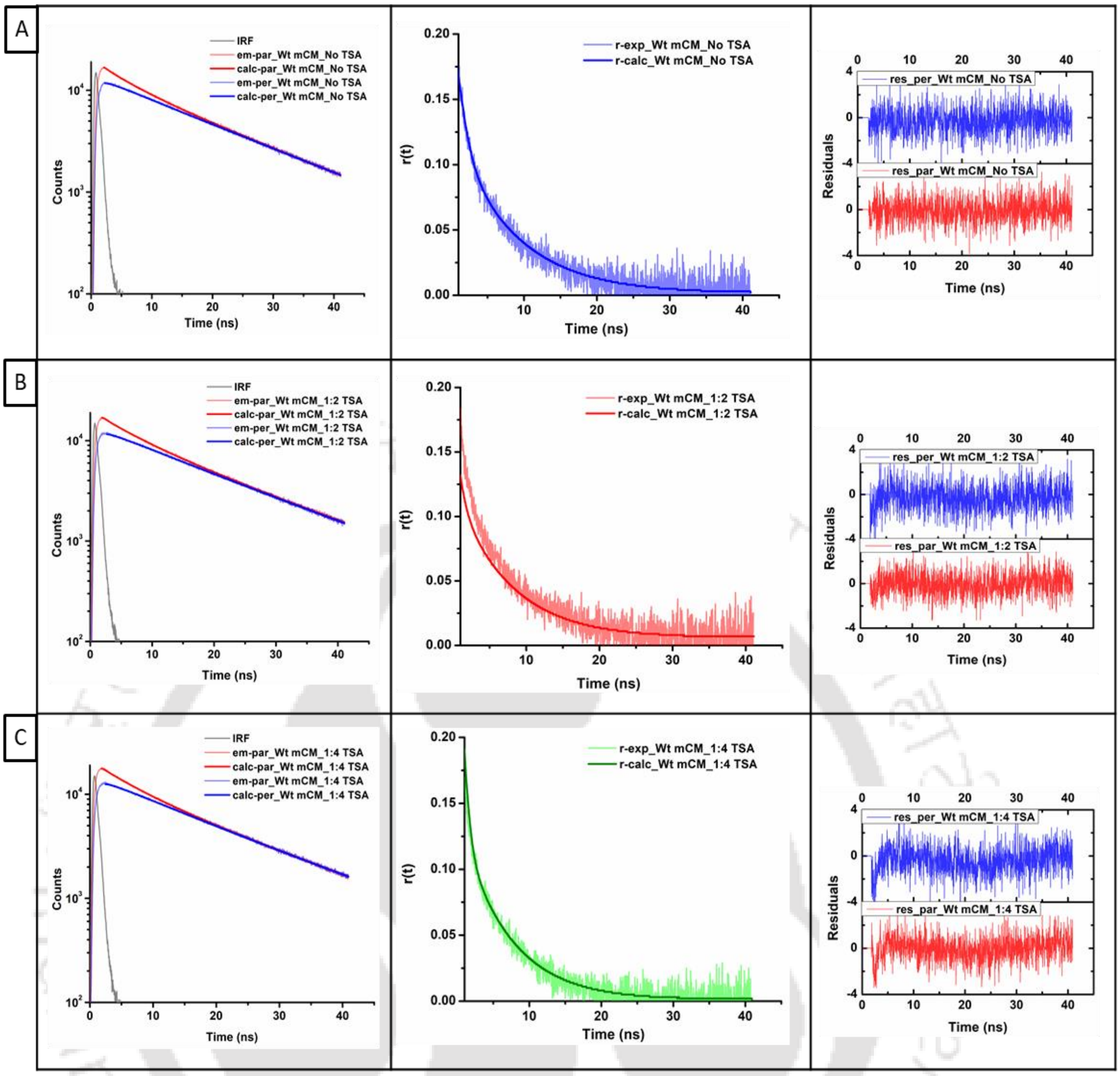


Figure 5.19. Dansyl anisotropy decay profile of Wt mCM in unbound and bound forms (with TSA). Residuals for the fitted data are shown in the middle panel. Decay plots, r(t) and residuals for fitted data of dansyl probe in Wt mCM with no added TSA (Top row, A), 1: 2 ratio TSA (Middle row, B) and 1: 4 ratio TSA (Bottom row, C) are shown.

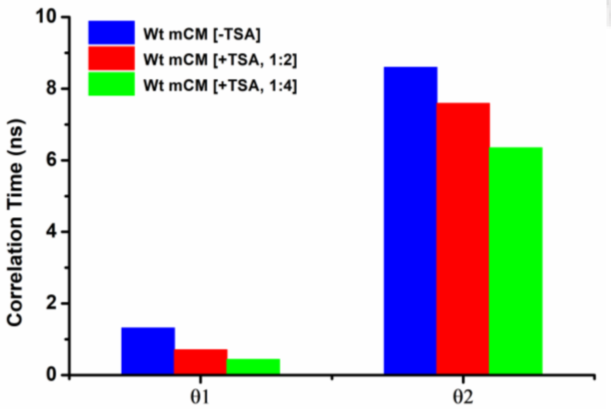


Figure 5.20. Histogram representing a comparison of the correlation time(s) in dansyl labeled Wt mCM in its unbound and TSA bound states. The fitted anisotropy decay parameters and std. dev. are listed in Table 5.10.

5.2.2. Time-resolved fluorescence anisotropy decay of dansyl probe in Mutant1

In case of the Mutant1, a fast rotational component of 1.2 ns with amplitude 0.37 ns and a slow rotational component of 10.3 ns with amplitude 0.63 was recorded after analysis of anisotropy decay profile of the dansyl probe (**Table 5.11, Figure 5.21**). Compared to the unbound Wt mCM, Mutant1 (unbound) displayed a longer time for the slow rotational component suggesting a less compact (larger) global structure of the protein after undergoing mutation (W24K). CD spectra (refer Chapter 4) of Wt mCM and the Mutant1 have also shown this change in global structure of mCM.

After TSA binding (60 µM and 120 µM), a more compact global structure of mCM in observed in Mutant 1 indicated by the decrease in the value of the slow rotational component to ~8.4 ns in its ligand bound form. But the compaction displayed in Mutant1 after TSA binding is visibly lesser than the Wt mCM. As in Wt mCM, ligand binding leads to a more flexible dansyl probe as supported by faster segmental motion in TSA bound state coupled with larger amplitude for short correlation time (**Table 5.11**). But the degree of flexibility in Mutant1 seems to be less in comparison to Wt mCM (**Table 5.11**).

Table 5.11. Anisotropy decay data of dansyl labeled Mutant1 (-/+ TSA) in PBS buffer pH 7.4

Mutant1	**r_0[a]**	**r_{ss}[b]**	**θ_1[c] (ns)**	**β_1[d]**	**θ_2[c] (ns)**	**β_2[d]**	**χ^2[e]**
No TSA	0.18	0.013	1.2 ± 0.06	0.37	10.3 ± 0.15	0.63	1.17
1:2 TSA	0.19	0.014	0.98 ± 0.03	0.49	8.25 ± 0.13	0.51	1.23
1:4 TSA	0.19	0.014	0.72 ± 0.02	0.52	8.47 ± 0.13	0.48	1.23

[a] initial anisotropy; [b] steady state anisotropy calculated from fit; [c] rotational correlation time(s); [d] fractional amplitude associated with the correlation time; [e] reduced χ^2 for the fit.

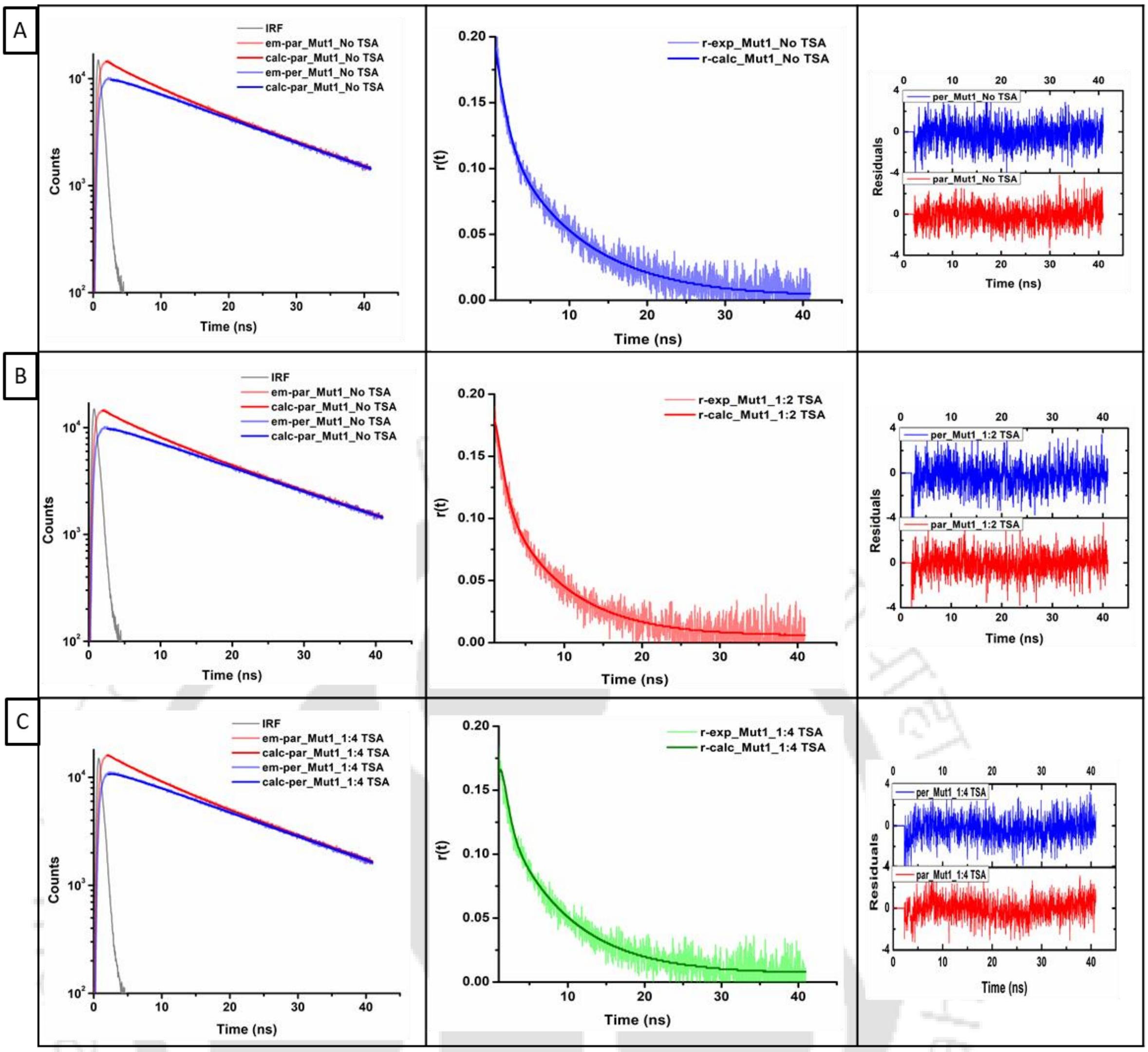


Figure 5.21. Dansyl anisotropy decay profile of Mutant1 in unbound and bound forms (with TSA). Residuals for the fitted data are shown in the middle panel. Decay plots, r(t) and residuals for fitted data of dansyl probe in Mutant1 with no added TSA (Top row, A), 1:2 ratio TSA (Middle row, B) and 1:20 ratio TSA (Bottom row, C) are shown.

A comparison of the correlation time(s) in dansyl labeled Mutant1 in its unbound and TSA bound states are shown in **Figure 5.22.**

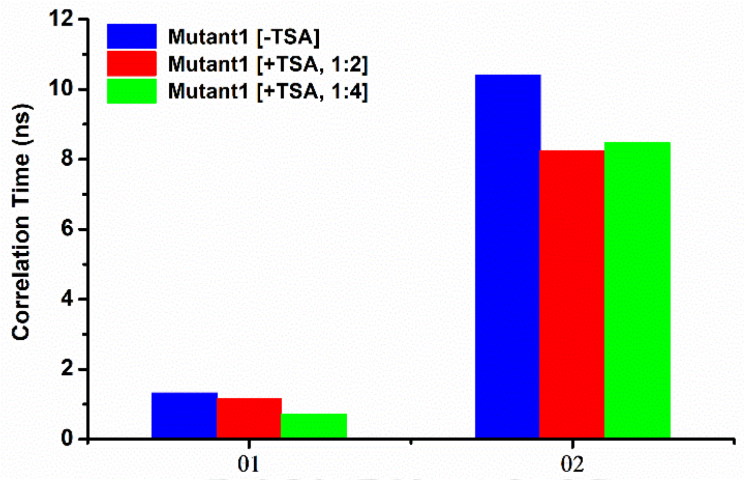


Figure 5.22. Histogram representing a comparison of the correlation time(s) in dansyl labeled Mutant1 in its unbound and TSA bound states. The corresponding amplitude and std dev. for each correlation time in the histogram in mentioned in Table 5.11

5.2.3 Time-resolved fluorescence anisotropy decay of tryptophan in Mutant2

The double mutant (Mutant2) lacks cysteine residue and therefore its tryptophan anisotropy decay was recorded in its unbound and TSA bound states (**Table 5.12, Figure 5.23**). A fast rotational component of 0.52 ns represented the single tryptophan probe (Trp 26) in a highly flexible loop region of mCM. Similar to analysed data obtained from dansyl probe in Mutant1, a slow rotational component of 10 ns was reported. This indicates that Mutant2 also retains a fairly less compact global structure than Wt mCM. Notably, a higher concentration of TSA (120 µM) was required in Mutant2 to attain a similar degree of compaction as in Mutant1. An excess concentration of 600 µM TSA incubated with Mutant2 displayed a slow rotational component of 7.5 ns, comparable to values obtained for Wt mCM with 60 µM TSA. Earlier lifetime fluorescence measurements also suggest that a higher concentration of TSA is required in the mutant variants, especially Mutant2 and Mutant3, to achieve noticeable changes in structure and dynamics of mCM. Another observation in Mutant2 is the absence of any change in the short rotational component after TSA binding implying the non-involvement of the flexible Trp loop region in the structural transitions associated with ligand binding.

An important thing to note here is that the tryptophan anisotropy decay profiles in all conditions for Mutant2 has an r_∞ component of ~0.06, the decay doesn't reach zero.

This is possibly due to the short lifetime of tryptophan (~2 ns) which gives a very small time window for tumbling, hence we cannot sample the long correlation time.

Table 5.12. Anisotropy decay data of tryptophan in Mutant2 (-/+ TSA) in PBS buffer pH 7.4

Mutant2	r_0[a]	r_{ss}[b]	θ_1[c] (ns)	β_1[d]	θ_2[c] (ns)	β_2[d]	χ^2[e]
No TSA	0.15	0.007	0.52 ± 0.02	0.86	10 ± 0.3	0.14	1.16
1:2 TSA	0.16	0.004	0.54 ± 0.02	0.91	9 ± 0.15	0.09	1.11
1:4 TSA	0.17	0.005	0.59 ± 0.015	0.90	8 ± 0.13	0.1	1.10
1:20 TSA	0.16	0.004	0.57 ± 0.018	0.87	7.5 ± 0.15	0.13	1.03

[a] initial anisotropy; [b] steady state anisotropy calculated from fit; [c] rotational correlation time(s); [d] fractional amplitude associated with the correlation time; [e] reduced χ^2 for the fit.

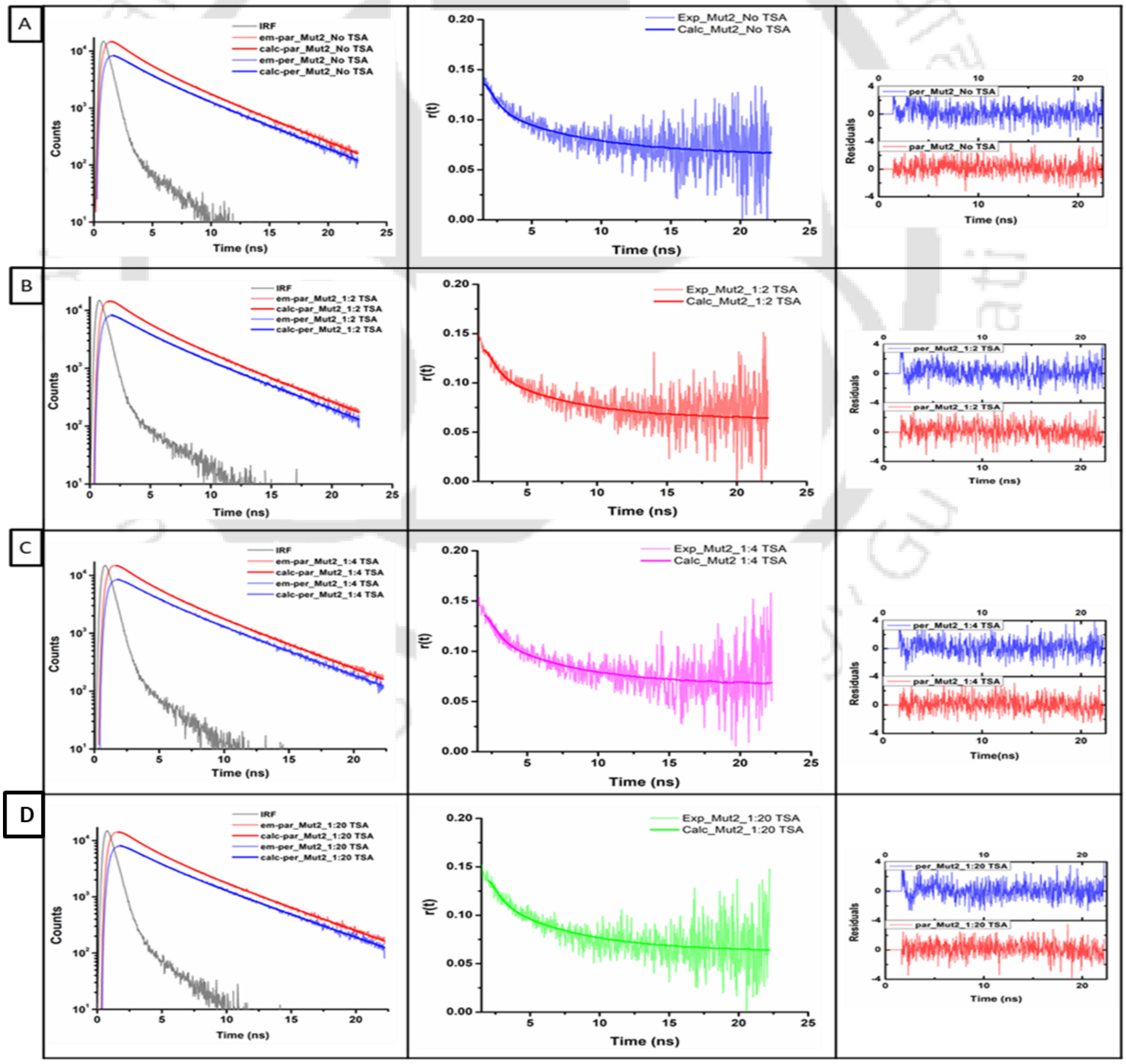


Figure 5.23. Dansyl anisotropy decay profile of Mutant2 in unbound and bound forms (with TSA). Decay plots, r(t) and residuals for fitted data of dansyl probe in Mutant3

with no added TSA (1st row, A), 1:2 ratio TSA (2nd row, B), 1:4 ratio TSA (3rd row, C) and 1:20 ratio TSA (4th row) are shown.

5.2.4 Time-resolved fluorescence anisotropy decay of dansyl probe in Mutant3

Analysis of the anisotropy decay profile of the dansyl probe in the triple mutant (Mutant3) revealed two correlation time, a short rotational component of 0.8 ns with an amplitude of 0.56 and a long rotational component of 10.93 with an amplitude of 0.44 (**Table 5.13, Figure 5.25**). The dansyl probe was placed at position Cys 6 in Mutant3 and hence provided information of a site different that the Wt mCM and Mutant1. In case of unbound Mutant3, the dansyl probe possess more flexibility than in Wt mCM and Mutant1 indicating a less hindered environment around Cys 6 residue than Cys 69 residue as evident by increased α_1 and faster ϕ_1 in presence of TSA. Moreover, Mutant3 displayed the least compact global structure amongst all mCM variants, inferred from its slow rotational component of ~11 ns; higher than Wt mCM and the other mutants.

After TSA binding, Mutant3 reported the minimal change in its global structure even after 600 µM TSA was used (**Table 5.13**). This result is complimentary to those obtained from lifetime FRET experiments (**Figure 5.11**) in Mutant3 reported earlier in this chapter (**Section 5.1.4**), which also indicates the aberrant nature of TSA binding in Mutant3. Report from CD spectra of Mutant3 showing a significant change in secondary structure compared to Wt mCM also points towards significant difference in overall global structure and hence its probable implications in TSA binding properties.

Table 5.13. Anisotropy decay data of dansyl labeled Mutant3 (-/+ TSA) in PBS buffer pH 7.4

Mutant3	r_0^a	r_{ss}^b	θ_1^c (ns)	β_1^d	θ_2^c (ns)	β_2^d	χ^{2e}
No TSA	0.15	0.013	0.8 ± 0.04	0.56	10.93 ± 0.16	0.44	1.19
1:2 TSA	0.16	0.012	0.7 ± 0.04	0.61	10.33 ± 0.26	0.39	1.19
1:20 TSA	0.16	0.013	0.61 ± 0.03	0.66	10.56 ± 0.22	0.34	1.21

[a] initial anisotropy; [b] steady state anisotropy calculated from fit; [c] rotational correlation time(s); [d] fractional amplitude associated with the correlation time; [e] reduced χ^2 for the fit.

A comparison of the correlation times in dansyl labelled Mutant3 in its unbound and TSA bound states are shown in **Figure 5.24**

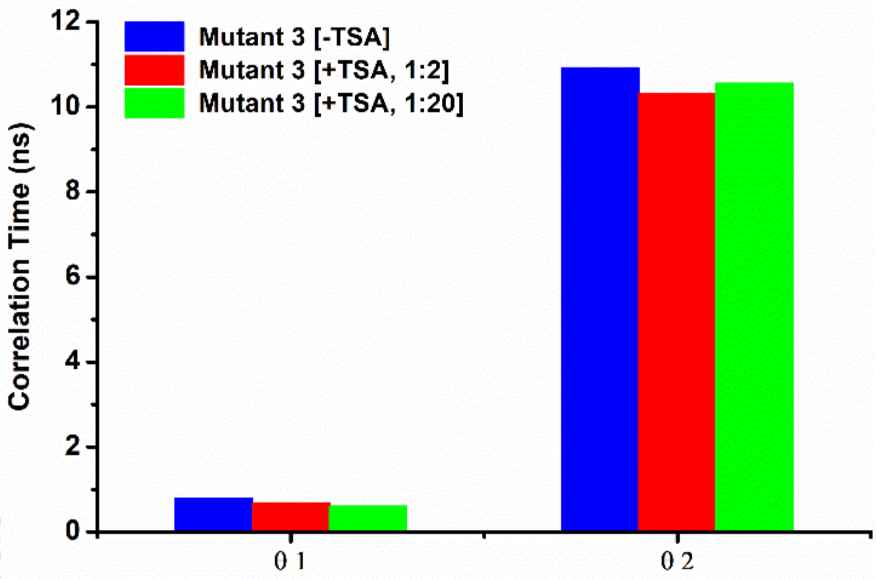


Figure 5.24. Histogram representing a comparison of the correlation time(s) in dansyl labeled Mutant3 in its unbound and TSA bound states. The corresponding amplitude and std. dev. for each correlation time in the histogram in mentioned in Table 5.13.

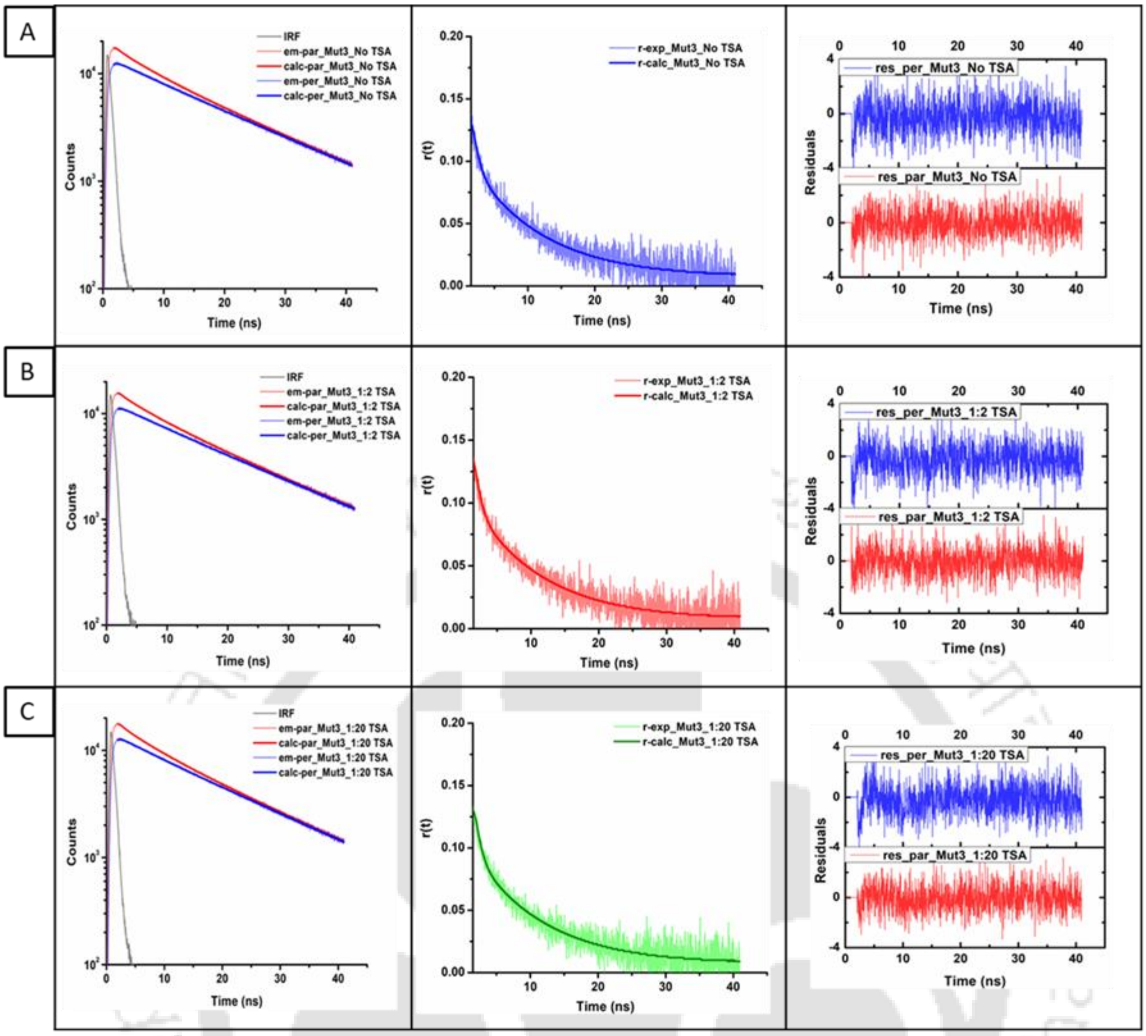


Figure 5.25. Dansyl anisotropy decay profile of Mutant3 in unbound and bound forms (with TSA). Decay plots, r(t) and residuals for fitted data of dansyl probe in Mutant3 with no added TSA (Top row, A), 1:2 ratio TSA (Middle row, B) and 1:20 ratio TSA (Bottom row, C) are shown.

5.3. Conclusions

- Fluorescence lifetime measurements of tryptophan in unlabeled form of mCM indicated that the lifetime of the two tryptophans in Wt mCM was significantly lower than in its mutant variants, supporting the existence of homoFRET between the closely placed tryptophan residues, as reported earlier in this thesis (Chapter 4).

- The single tryptophan mutant, Mutant1, provided vital information on structural transitions in mCM after TSA binding. A significant decrease in FRET distance from 23 Å in unbound form to 16 Å in ligand bound form was observed indicating tightening to the protein structure around the ligand. This result is in agreement with NMR data of mCM-TSA which reported a Trp26-Cys69 distance of ~15.6 Å

- Mutant3 reported a decrease in FRET distance from ~29.3 Å to ~25.7 Å between Trp26-Cys6 which also clearly suggested the overall tightening of mCM structure after TSA binding. A high concentration of TSA required to observe this effect in Mutant2 indicates towards its structural and functional perturbations due to the point mutations.

- Time-resolved fluorescence anisotropy decay studies of the dansyl probe attached to Cysteine residues in Wt mCM, Mutant1 and Mutant 3 revealed that the mutants were less compact in structure than the wild form but gained compaction after binding to the TSA.

- As with the lifetime fluorescence measurements, time-resolved anisotropy decay profile of Mutant3 also reported the aberrant nature of TSA binding in this mutant.

- Tryptophan anisotropy decay profile of Mutant2 reported a global structure similar to Mutant1 which underwent compaction also comparable to the Mutant1 after TSA binding even though a higher concentration of TSA was required, indicating structural and functional alternations in Mutant2. However, the presence of an r_∞ component means that long correlation time for Mutant2 could not be sampled properly in the short tryptophan lifetime window and hence, the data may not be accurately recorded.

- MEM distributions mostly overlay on the discrete lifetimes obtained by non-linear least square analysis. The width of a single distribution peak can indicate the uncertainty in that lifetime, implying insufficient information (viz. limited time resolution) in the intensity decay to resolve the distribution any further than already done.

Chapter-6

Assay of mCM enzyme activity

6.1. Enzyme activity assay of Wt mCM

To measure the enzymatic activity of mCM, the disappearance of its substrate (Chorismic acid) was monitored spectroscopically at 274 nm (ε_{274} = 2630 $M^{-1}cm^{-1}$). Readings were recorded in a UV visible spectrophotometer (make: Perkin Elmer) at every 1 s interval for 300 s. 150 nM of mCM enzyme in PBS buffer pH 7.2 supplemented with bovine serum albumin (0.1 mg/ml) was used for the assay. A varying substrate concentration of 50-400 µM was utilized for the enzymatic reaction. The protocol used was as previously described by Hilvert group, ETH Zurich (Vamvaca *et al.*, 2004) and was repeated in case of all the mutants.

Figure 6.2 (*Left*) represented the absorbance versus time plot of chorismic acid in presence of Wt mCM. The data obtained is extrapolated in the Eadie-Hofstee plot to obtain the K_m (concentration of substrate when ½ V_{max} is reached), V_{max} (maximum velocity) and k_{cat} (turnover number) values. The Eadie-Hofstee plot is based on the transformation of the Michaelis-Menten equation to give the equation of a straight line. The initial velocity (v) value is plotted along the y-axis and v/[S], where [S] is the substrate concentration, along the x-axis. The slope of the best-fit line is equal to $-K_m$, the y-intercept is equal to V_{max}, and the x-intercept is equal to V_{max}/K_m as shown in **Figure 6.1**

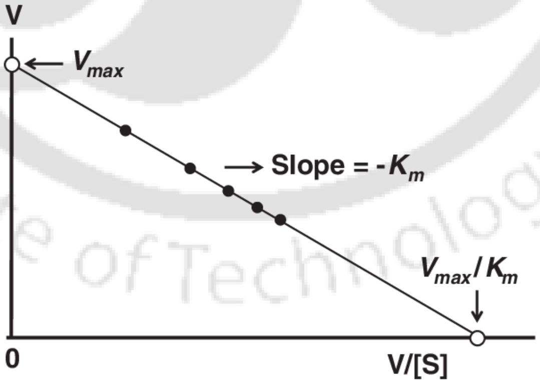


Figure 6.1. Eadie-Hofstee plot for enzyme kinetics. The v value is plotted along the y-axis and v/[S] along the x-axis. The slope of the best-fit line is equal to $-K_m$, the y-intercept is equal to V_{max}, and the x-intercept is equal to V_{max}/K_m.

Figure 6. 2 (*Right*) shows the Eadie-Hofstee plot for Wt mCM from which K_m (165 ± 46 µM) and k_{cat} (7.7 ± 0.17 s^{-1}) values obtained are nearly comparable to those earlier by

the Hilvert group (K_m 180 μM, k_{cat} 3.2 s^{-1}) (Vamvaca *et al.*, 2004)). These kinetic parameters (K_m, V_{max} and k_{cat}) were obtained for all the mutants and compared it with the Wt mCM for a quantitative measure of functional aberrations previously observed in our fluorescence and CD studies.

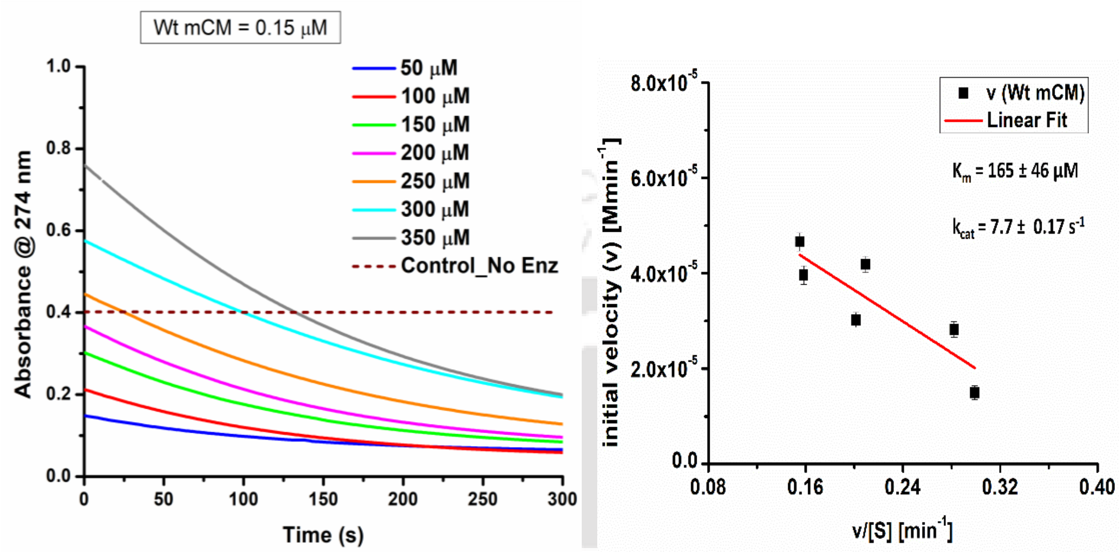


Figure 6.2. (*Left*) The absorbance versus time plot for the utilization of the substrate chorismic acid (50 μM-350 μM) by Wt mCM (150 nM) in PBS buffer pH 7.2 with BSA (1mg/ml). The control sample (maroon dashed line) contains chorismic acid (150 μM) without the enzyme in the sample mixture. (*Right*) Eadie Hofstee plot with a linear fit (red) giving the values for the slope ($-K_m$, 165 ± 46μM), y-intercept (V_{max}, $6.93*10^{-5} \pm 0.00001$ $Mmin^{-1}$) and R-square (0.87).

The turnover number (kcat) representing the maximal number of chemical conversions of substrate molecules per second for a single catalytic site and a given enzyme concentration is calculated from the following equation:

$$k_{cat} = \frac{V_{max}}{[E_t]} \tag{6.1}$$

where [Et] is the total enzyme concentration.

6.2. Enzyme activity assay of Mutant1

The single tryptophan mutant, Mutant1, retained much of the enzymatic activity of its wild-type in assays (**Figure 6.3**). Since the mutated tryptophan residue (Trp 24) is located in a flexible loop region linking two helices, the change from tryptophan to lysine residue seems to not significantly affect the structural and functional attributes of mCM.

The slight changes in K_m (299 ± 91 µM) and k_{cat} (11.5 ± 4.5 s^{-1}) values correlated with minor secondary structural changes in our CD measurements (**Chapter 4**, **Figure 4.6**.). Earlier report by the Hilvert group (Vamvaca *et al.*, 2004)) suggested the absence of any negative impact on mCM functionality after the Trp 24 to Lys 24 mutation.

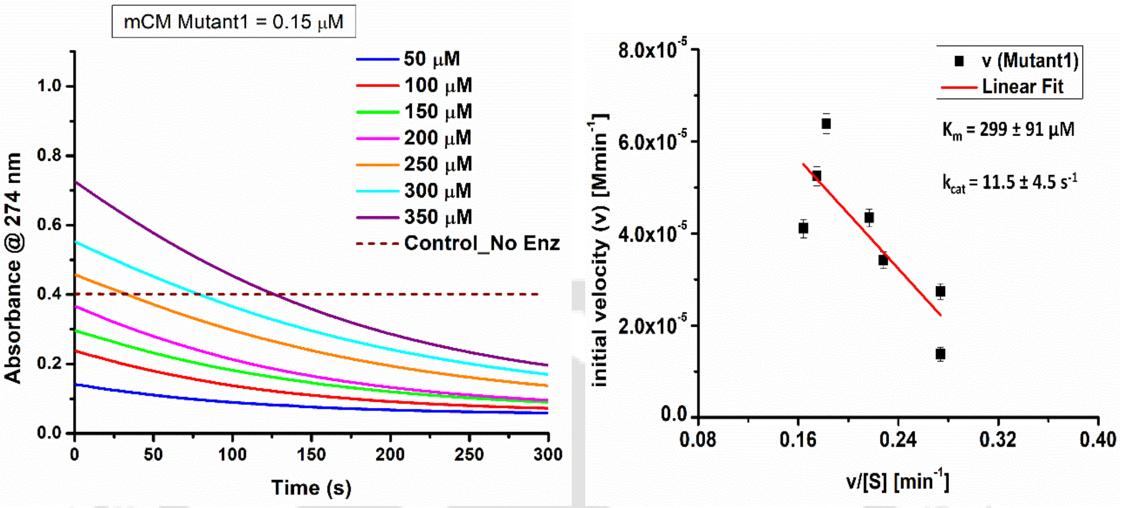


Figure 6.3. (*Left*) The absorbance versus time plot for the utilization of the substrate chorismic acid (50 µM-350 µM) by Mutant1 (150 nM) in PBS buffer pH 7.2 with BSA (1mg/ml). The control sample (maroon dashed line) contains chorismic acid (150 µM) without the enzyme in the sample mixture. (*Right*) Eadie-Hofstee plot with a linear fit (red) giving the values for the slope (-K_m, 299 ± 91 µM), y-intercept (V_{max}, $1.041*10^{-4}$ ± 0.00002 $Mmin^{-1}$) and R-square (0.82).

6.3. Enzyme activity assay of Mutant2

The double mutant, Mutant2, displayed a significant decline in enzyme activity in our assays. Our CD measurements (**Chapter 4**, **Figure 4.6**.) and ANS assays (**Chapter 4**, **Figure 4.3**) also indicated significant deviation in secondary structural elements of Mutant2 from Wt mCM. These structural differences in Mutant2 due to the Cys 69 to Ala 69 mutation (a cysteine-less mutant) may be the root cause for functional loss of Mutant2. With 100 nM of the enzyme, negligible change in absorbance of chorismic acid was observed in **Figure 6.4** (*Left*), while addition of an excess of enzyme to the reaction mixture showed significant decrease in absorbance values (**Figure 6.4, *Right***). From **Figure 6.5**, K_m (603 ± 104 µM) and k_{cat} (0.68 ± 0.46 s^{-1}) values were obtained in presence of an excess of the Mutant2 enzyme (33 fold, 5 µM).

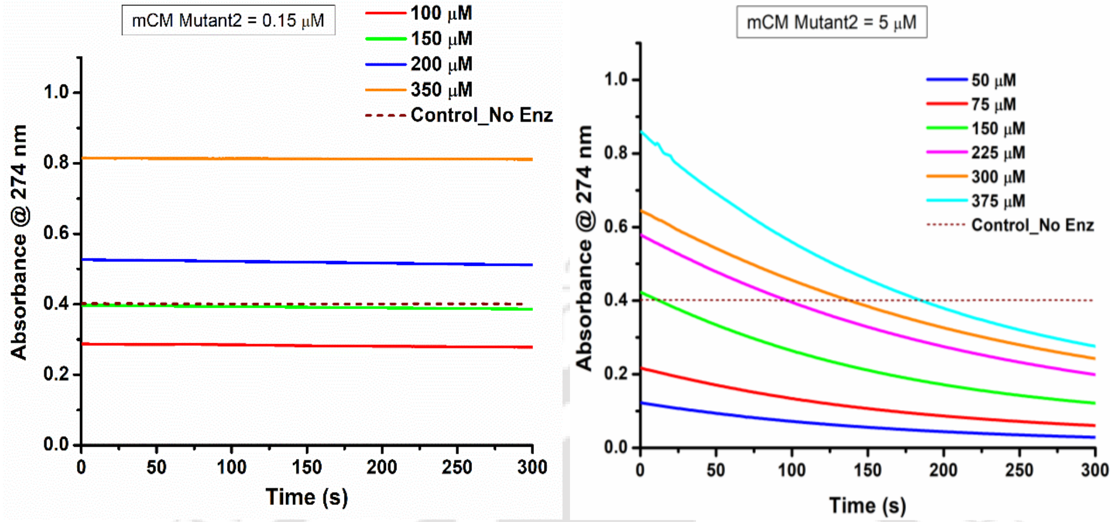


Figure 6.4. (*Left*) The absorbance versus time plot for the utilization of the substrate chorismic acid (100 µM-350 µM) by Mutant2 (150 nM) in PBS buffer pH 7.2 with BSA (1mg/ml). The control sample (maroon dashed line) contains chorismic acid (150 µM) without the enzyme in the sample mixture. (*Right*) The absorbance versus time plot for the utilization of the substrate chorismic acid (50 µM-350 µM) by Mutant2 (5 µM).

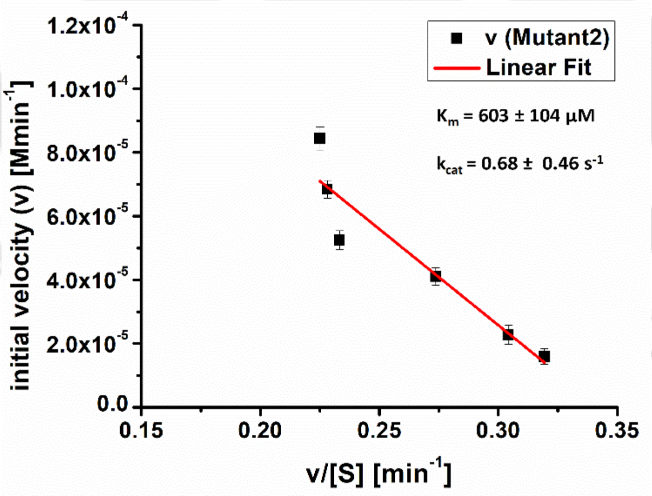


Figure 6.5. Eadie-Hofstee plot for Mutant2 (5 µM) with a linear fit (red) giving the values for the slope (-K_m, 603 ± 104 µM), y-intercept (V_{max}, $2.067*10^{-4} \pm 0.00002$ $Mmin^{-1}$) and R-square (0.94).

6.4. Enzyme activity assay of Mutant3

The triple mutant, Mutant3 also displayed an activity level much lower than the Wt mCM, but surprisingly it was more active than Mutant2.

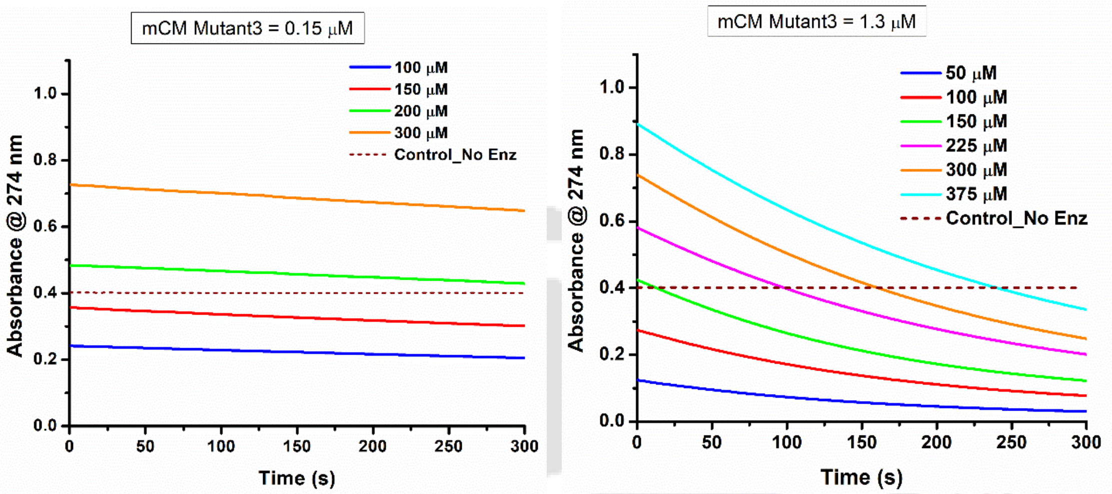


Figure 6.6. (*Left*) The absorbance versus time plot for the utilization of the substrate chorismic acid (100 μM-300 μM) by Mutant3 (150 nM) in PBS buffer pH 7.2 with BSA (1mg/ml). The control sample (maroon dashed line) contains chorismic acid (150 μM) without the enzyme in the sample mixture. (*Right*) The absorbance versus time plot for the utilization of the substrate chorismic acid (50 μM-350 μM) by Mutant3 (1.3 μM).

In case of Mutant3, the absorbance versus time plot for chorismic acid showed slight decrease in absorbance in presence of 150 nM of the enzyme (**Figure 6.6, *Left***). With an excess concentration of Mutant3 (1.3 μM), a substantial decrease in absorbance of chorismic acid was observed (**Figure 6.6, *Right***), comparable to Wt mCM. This was used to obtain a K_m (295 ± 46 μM) and k_{cat} (1.45 ± 0.18 s^{-1}) values through Eadie-Hofstee plot (**Figure 6.7**)

The re-introduction of Cysteine at position 6 (Ala6 to Cys6 mutation) in this mutant interestingly seemed to recover some degree of functionality in the enzyme. Our CD measurements also revealed a recovery of α-helicity in Mutant3 when compared with the cysteine-less Mutant2. It seems that the single cysteine residue (Cys 69) plays a crucial role in the structural and functional integrity of mCM.

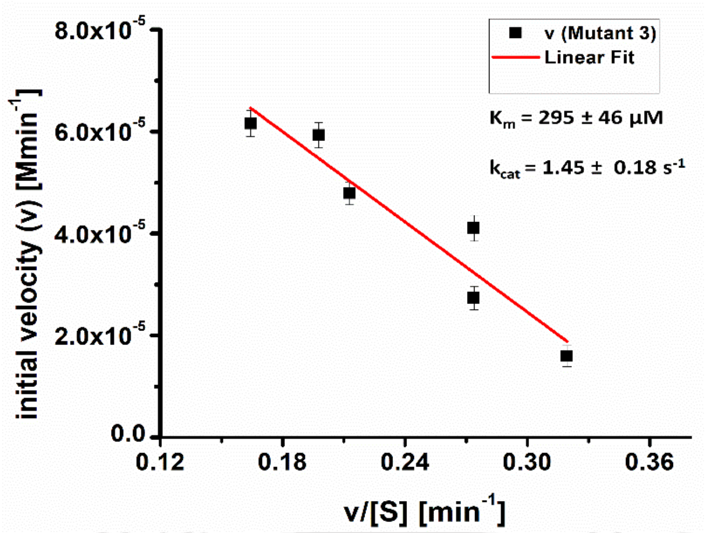


Figure 6.7. Eadie-Hofstee plot for Mutant3 (1.3 µM) with a linear fit (red) giving the values for the slope (-K_m, 295 ± 46 µM), y-intercept (V_{max}, $1.131*10^{-4} \pm 0.00001$ $Mmin^{-1}$) and R-square (0.95).

Table 6. Comparison of the kinetic parameters of Wt mCM with its mutants.

Sample	K_m [µM]	k_{cat} [s^{-1}]	Catalytic Efficiency k_{cat}/K_m [$M^{-1}s^{-1}$]	Percentage activity
Wt mCM	165 ± 46	7.7 ± 0.17	$4.7\times10^{4} \pm 0.007$	100
Mutant 1	299 ± 91	11.5 ± 4.5	$3.8\times10^{4} \pm 5.06$	85.1
Mutant 2	603 ± 104	0.68 ± 0.46	$1.1\times10^{3} \pm 0.053$	2.5
Mutant 3	295 ± 46	1.45 ± 0.18	$4.9\times10^{3} \pm 0.008$	10.9

The percentage drop in enzyme activity was calculated from **Table 6** as 15% for Mutant1, a drooping 97% for Mutant2 and 89% for Mutant3.

6.5. Conclusions

- An enzymatically active Wt mCM was expressed and purified and the enzyme kinetic parameters were quantified similar to earlier studies.
- Mutant1 retained the enzyme activity to a large extent whereas Mutant2 and Mutant3 showed a 97% and 89% dip in activity respectively.
- The highest change was recorded for the cysteine lacking double mutant, Mutant2, which inferred the crucial role of the single cysteine residue in mCM function.

CHAPTER-7

Discussion

The classical view in proteomics and structural biology that a fixed structure enables a protein to carry out its designated role(s) has experienced a shift in paradigm with emerging evidence suggesting that a large number of proteins can function even in the absence of a stable, folded structure (Uversky 2019, Uversky 2019). It is estimated that close to 40% of the proteins encoded by the human genome are intrinsically disordered or contain disordered regions of significant (>30 amino acids) length (van der Lee, Buljan et al. 2014). The IDPs exist as interconverting conformational ensembles, without any fixed three-dimensional structure. Intrinsically disordered enzymes, such as mCM, presents a rare and compelling case to improve our understanding of how enzymes function. Information from detailed investigations into the functional transitions adopted by this highly flexible yet efficient enzyme can aid in expanding the realm of functionally viable engineered enzymes for various pharmaceutical and industrial uses.

We set out to investigate the structural plasticity mechanisms that modulate catalysis by the molten globule mCM. Capturing the structural transitions in mCM during catalysis will provide us valuable insight into our understanding of inner workings of the disordered enzymes. Understanding how enzymes respond to mutations can provide valuable insight into the structure-function relationships (Peracchi 2001, Toscano, Woycechowsky et al. 2007).We then probed the effects of rationally designed mutations on segmental dynamics, conformational adaptation, disorder-to-order transition and catalytic efficiency to dissect out the determinants and influencers of catalysis in molten globule enzymes. Our model enzyme was amenable to the sensitive fluorescence spectroscopy tools at our disposal. We generated three mutants namely Mutant1 (W24K), Mutant2 (W24K/C69A) and Mutant3 (W24K/C69A/A6C). The rationale behind the selection of these designer mutants has been described in detail in the introduction section.

In this thesis, extensive studies were performed to capture both minor and major changes in the tertiary conformation of mCM and correlate it with its functions. Through time-resolved FRET experiments, we could sensitively measure and quantify (in terms of distance) local segmental dynamics and global transitions during binding of the ligand (TSA) to mCM. The time-resolved anisotropy studies supported the concept of a global collapse (ligand-induced global compaction) of the mCM structure after ligand binding.

The engineered molten globule enzyme with its relatively small size and simple secondary structure, can serve as a good model system for us to experimentally study and subsequently exert better control over enzyme catalysed reactions as well as disease causing disordered proteins. The molten globule enzyme can undergo efficient catalysis probably due to the ligand-induced global compaction that introduces conformational order thereby facilitating catalysis. The ordering of the global conformation observed upon binding to the ligand could indicate that folding can be coupled to catalysis with minimal energetic penalty. It was found that the cysteine mutants (Mutant2 and Mutant3) possessed a less compact global tertiary conformation even after ligand binding. These site-specific mutations may have impaired the ability of mCM to undergo effective ligand induced disorder-to-order transitions, resulting in a decrease in global structural ordering when compared to Wt mCM. This is the probably affecting catalytic activity as the mutants show drastically lower enzyme activity. Hence, the sensitivity of this disordered protein to certain point mutations is one of the key highlights of the thesis as shown in **Figure 7. Mutant3** (W24K/C69A/A6C) showed some degree of compensatory behavior both in alpha helicity, chemical stability, and enzyme activity when compared to **mutant2** (W24K/ C69A) indicating local structural plasticity playing a plausible role.

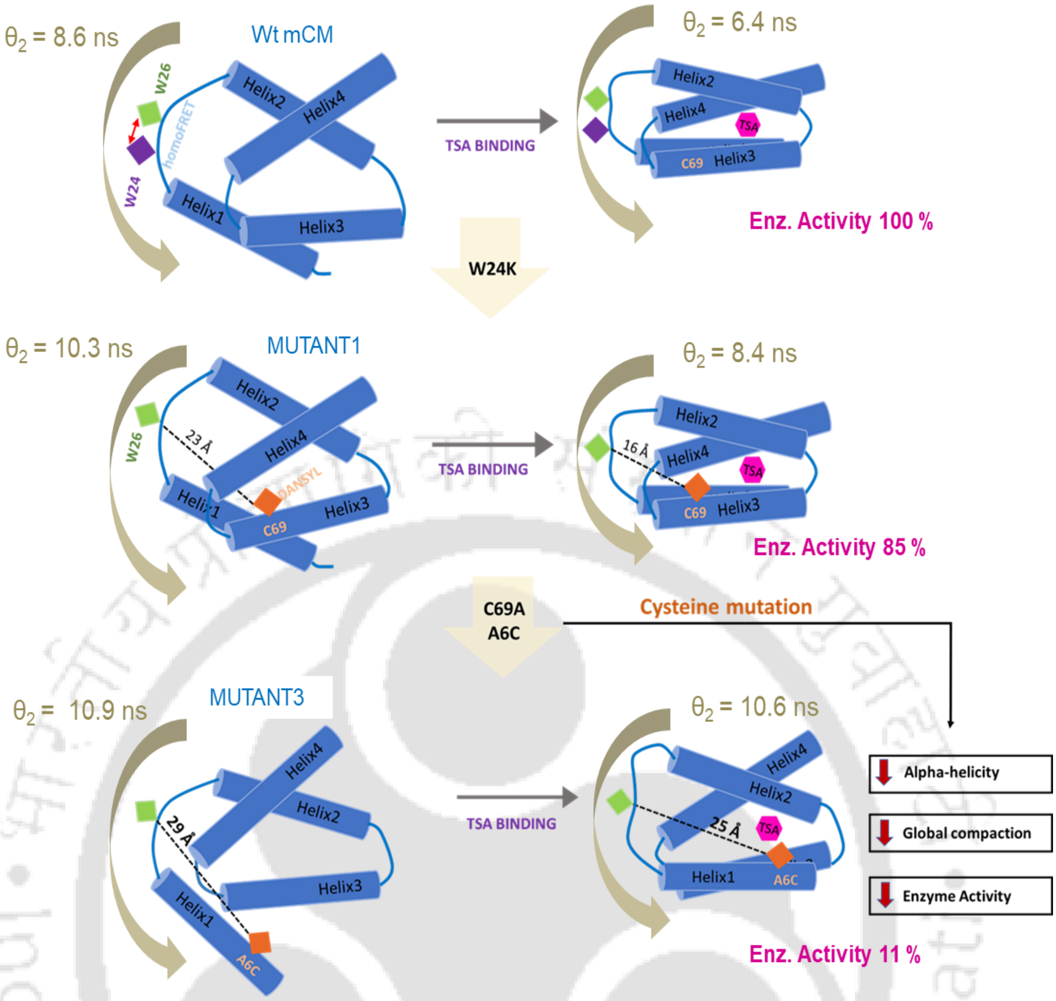


Figure 7. Diagram highlighting some of the key findings in this thesis work.

A significant loss in enzyme activity and change in secondary structure in the non-active site cysteine mutants strongly suggests a role of the single cysteine (Cys69) in maintaining the secondary structure and functional integrity of the mCM enzyme. Interestingly, ANS fluorescence assay showed that TSA can bind to the core hydrophobic region and displace ANS even in the cysteine mutants, Mutant2 and Mutant3. Thus, the cysteine mutations might not have affected ligand binding to a considerable extent. Instead, our results indicate that the hydrophobic collapse of the entire mCM structure following ligand binding might be the process severely affected by the mutations. In this regard, the free single cysteine residue (C69, helix 3) may be an important element in the network driving function-associated hydrophobic collapse of mCM. An important property of mCM to consider, as an IDP, is its inherent flexibility that speeds up recognition of its interacting ligand (TSA) by almost threefold compared to the ordered

CM dimer (Vamcava et al. 2008). IDPs have been known to form weak long-range protein-ligand interactions (owing to the ‘fly-casting mechanism’ discussed earlier in Chapter1) which strengthen as the protein accommodates its target, thereby accelerating molecular recognition. mCM probably makes initial contacts with its ligand through such long-range interactions where the non-active site residue Cys 69 has a role to play. Loss of direct contact with the ligand or a loss of α-helicity in the vicinity of Cys 69 (helix-3) can compromise initial interaction with the ligand, severely affecting binding and catalysis.

Many IDPs show context-dependent modulation of protein folding landscape inkling towards the fact that the primary amino-acid sequences partially codes for the proteins native folds while the rest is completed upon interaction with binding partners (Moosa, Ferreon et al. 2015). This could be the mechanism by which IDPs exhibit enormous structural plasticity and interaction promiscuity. The high context dependence of the effects of mutations in enzymology and IDP research has been actively investigated. Interestingly, in an earlier study investigating the effects of point mutations at the active site of mCM reported that the molten globule is relatively more resilient to mutation when compared to analogous mutation of the natural thermostable chorismate mutase (Woycechowsky, Choutko et al. 2008). The activity difference between the two enzymes stemmed from context-dependent, subtle changes in fine structure of active site rather than loss of secondary structural integrity or thermodynamic stability. In our study, a non-active site mutation showed dramatic effects in enzyme activity due to changes in secondary structure, ligand-induced global compaction and chemical denaturation states. This is highly intriguing due to the observation made that perturbation of single residues at the non-active sites can have more dramatic effects that perturbation at the active site as observed in case of mCM. This is quite in contrast to what one would anticipate of the consequences of mutations at active vs non-active sites. The thermodynamics of such perturbation is a subject matter of further investigation. The structural diversity of protein states can be harnessed for regulatory purposes.

In case of IDPs, stochastic conformational switching (Choi, McCann et al. 2011), crowding-induced compaction (Banks, Qin et al. 2018, Fonin, Darling et al. 2018), heat-induced compaction (Langridge, Tarver et al. 2014), length-induced compaction (Uversky, Santambrogio et al. 2012) could make significant role in its mechanism of action. It was earlier observed by isothermal calorimetry that in case of mCM the structural plasticity may strongly perturb the thermodynamics and kinetics of transition-state recognition while negligibly affecting catalytic efficiency ((Vamvaca, Jelesarov et al. 2008). Interestingly, we find that integrity of secondary structure, segmental dynamics and ligand-induced global compaction is tightly linked to the mCM mediated enzyme catalysis.

The IDPs are structurally and functionally heterogenous and can be considered as important part of the structure-function continuum. In case of disordered enzymes, the disorder could be necessary for substrate recognition, binding and promiscuity. Unravelling the “codes” of the enzyme-ligand recognition and the regulating principles that modulate catalysis will help in the design of enzymes with tailored properties and significantly aid in our understanding of the biophysics of intrinsically disordered proteins.

References

Al-Ali H, Rieger ME, Seldeen KL, Harris TK, Farooq A, and Briegel KJ. 2010. "Biophysical Characterization Reveals Structural Disorder in the Developmental Transcriptional Regulator LBH." *Biochemical and Biophysical Research Communications* 391 (1): 1104–9.

Ansari MZ, Kumar A, and Ahari D.2018. "Protein charge transfer absorption spectra : an intrinsic probe to monitor structural and oligomeric transitions in proteins". *Farady Discussions* 207: 91–113.

Aumüller T, and Fischer G. 2008. "Bioactivity of Folding Intermediates Studied by the Recovery of Enzymatic Activity during Refolding." *Journal of Molecular Biology* 376(5):1478-92

Aviles FJ, Chapman GE, Kneale GG, Robinson CC, and Bradbury EM. 1978. "The Conformation of Histone H5." *European Journal of Biochemistry* 88 (2): 363–71.

Baker JM, Hudson RP, Kanelis V, Choy WY, Thibodeau PH, Thomas PJ, and Forman-Kay JD. 2007. "CFTR Regulatory Region Interacts with NBD1 Predominantly via Multiple Transient Helices." *Nature Structural & Molecular Biology* 14(8):738-45.

Baldwin, R. L. and B. H. Zimm (2000). "Are denatured proteins ever random coils?" *Proc Natl Acad Sci U S A* **97**(23): 12391-12392.

Banks, A., S. Qin, K. L. Weiss, C. B. Stanley and H. X. Zhou (2018). "Intrinsically Disordered Protein Exhibits Both Compaction and Expansion under Macromolecular Crowding." *Biophys J* **114**(5): 1067-1079.

Bartlett P, and Johnson C. 1985. "An Inhibitor of Chorismate Mutase Resembling the Transition-State Conformation." *J. Am. Chem. Soc.* 107**:**7792-7793.

Belford GG, Belford RL, and Weber G. 1972. "Dynamics of Fluorescence Polarization in Macromolecules." *Proc. Natl. Acad. Sci.* 69:1392-1393.

Bemporad F, Gsponer J, Hopearuoho HI, Plakoutsi G, Stati G, Stefani M, Taddei N, Vendruscolo M, and Chiti F. 2008. "Biological Function in a Non-Native Partially Folded State of a Protein." *The EMBO Journal* 27(10):1525-35.

Benkovic SJ, and Schiffer SH. 2003. "A Perspective on Enzyme Catalysis." *Science (New York, N.Y.).* 301(5637):1196-202.

Betz SF, Raleigh DP, and DeGrado WF. 1993. "De Novo Protein Design: From Molten Globules to Native-like States: Current Opinion in Structural Biology 1993, 3:601–610." *Current Opinion in Structural Biology* 3 (4): 601–10.

Bevington PR, and Robinson KD (1992): Data reduction and error analysis for the physical sciences. *McGraw-Hill, Inc.*

Bhatia S, Krishnamoorthy G, and Udgaonkar JB. 2018. "Site-Specific Time-Resolved FRET Reveals Local Variations in the Unfolding Mechanism in an Apparently Two-State Protein Unfolding Transition." *Physical Chemistry Chemical Physics* 20 (5): 3216–32.

Borg M, Mittag T, Pawson T, Tyers M, Forman-Kay JD, and Chan HS. 2007.

"Polyelectrostatic Interactions of Disordered Ligands Suggest a Physical Basis for Ultrasensitivity." *Proceedings of the National Academy of Sciences* 104 (23): 9650 LP-9655.

Bornhorst JA, and Falke JJ. 2000. "Purification of Proteins Using Polyhistidine Affinity Tags." *Methods in Enzymology* 326: 245–54.

Bracken C. 2001. "NMR Spin Relaxation Methods for Characterization of Disorder and Folding in Proteins." *Journal of Molecular Graphics and Modelling* 19 (1): 3–12.

Brown CJ, Takayama S, Campen AM, Vise P, Marshall TW, Oldfield CJ, Williams CJ, and Dunker AK. 2002. "Evolutionary Rate Heterogeneity in Proteins with Long Disordered Regions ." *Journal of Molecular Evolution* 55 (1): 104–10.

Bürgi, Jérôme, Bin Xue, Vladimir N Uversky, and F Gisou van der Goot. 2016. "Intrinsic Disorder in Transmembrane Proteins: Roles in Signaling and Topology Prediction." *PLOS ONE* 11 (7): e0158594.

Cheng Y, LeGall T, Oldfield CJ, Dunker AK, and Uversky VN. 2006. "Abundance of Intrinsic Disorder in Protein Associated with Cardiovascular Disease." *Biochemistry* 45 (35): 10448–60.

Choi, U. B., J. J. McCann, K. R. Weninger and M. E. Bowen (2011). "Beyond the random coil: stochastic conformational switching in intrinsically disordered proteins." *Structure* **19**(4): 566-576.

Choy WY, and Forman-Kay JD. 2001. "Calculation of Ensembles of Structures Representing the Unfolded State of an SH3 Domain." *Journal of Molecular Biology* 308 (5): 1011–32.

Crick SL, Jayaraman M, Frieden C, Wetzel R, and Pappu RV. 2006. "Fluorescence Correlation Spectroscopy Shows That Monomeric Polyglutamine Molecules Form Collapsed Structures in Aqueous Solutions." *Proceedings of the National Academy of Sciences* 103 (45): 16764 LP-16769.

Daughdrill GW, Pielak GJ, Uversky VN, Cortese MS, and Dunker AK. 2005. "Natively Disordered Proteins." *Protein Folding Handbook*. Wiley Online Books.

DeForte S and Uversky VN. 2017. "Not an Exception to the Rule: The Functional Significance of Intrinsically Disordered Protein Regions in Enzymes." *Molecular BioSystems* 13 (3): 463–69.

DeGrado W F. 1993. "Catalytic Molten Globules." *Nature* 365: 488–489.

Demarest SJ, Boice JA, Fairman R, and Raleigh DP. 1999. "Defining the Core Structure of the α-Lactalbumin Molten Globule." *Journal of Molecular Biology* 294 (1): 213–21.

Dicko, C., M. R. Hicks, T. R. Dafforn, F. Vollrath, A. Rodger and S. V. Hoffmann (2008). "Breaking the 200 nm limit for routine flow linear dichroism measurements using UV synchrotron radiation." *Biophys J* **95**(12): 5974-5977.

Dolgikh DA, Gilmanshin RI, Brazhnikov EV, Bychkova VE, Semisotnov GV, Venyaminov SY, and Ptitsyn OB. 1981. "α-Lactalbumin: Compact State with Fluctuating Tertiary Structure?" *FEBS Letters* 136 (2): 311–15.

Dosztányi Z, Chen J, Dunker AK, Simon I, and Tompa P. 2006. “Disorder and Sequence Repeats in Hub Proteins and Their Implications for Network Evolution.” *Journal of Proteome Research* 5 (11): 2985–95.

Dougan L, Li J, Badilla CL, Berne BJ, and Fernandez JM. 2009. “Single Homopolypeptide Chains Collapse into Mechanically Rigid Conformations.” *Proceedings of the National Academy of Sciences* 106 (31): 12605 LP-12610.

Dunker AK, Oldfield CJ, Meng J, Romero P, Yang JY, Chen JW, Vacic V, Obradovic Z, and Uversky VN. 2008. “The Unfoldomics Decade: An Update on Intrinsically Disordered Proteins.” *BMC Genomics* 9 (SUPPL. 2): 1–26.

Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, et al. 2001. “Intrinsically Disordered Protein.” *Journal of Molecular Graphics and Modelling* 19 (1): 26–59.

Dunker AK, Garner E, Guilliot S, Romero P, Albrecht K, Hart J, Obradovic Z, Kissinger C, and Villafranca JE. 1998. “Protein Disorder and the Evolution of Molecular Recognition: Theory, Predictions and Observations.” *Pacific Symposium on Biocomputing* 473–84.

Dunker AK, Obradovic Z, Romero P, Garner E, and Brown C. 2000. "Intrinsic Protein Disorder in Complete Genomes." *Workshop on Genome Informatics*. Vol. 11.

Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, and Obradović Z. 2002. “Intrinsic Disorder and Protein Function.” *Biochemistry* 41 (21): 6573–82.

Dunker AK, Brown CJ, and Obradovic J. 2002. “Identification and Functions of Usefully Disordered Proteins.” *Advances in Protein Chemistry* 62:25–49.

Dunker AK, Cortese MS, Romero P, Iakoucheva LM, and Uversky VN. 2005. “Flexible Nets.” *The FEBS Journal* 272 (20): 5129–48.

Dunker AK, and Obradovic Z . 2001. “The Protein Trinity—linking Function and Disorder.” *Nature Biotechnology* 19(9):805-6.

Dunker AK, Oldfield CJ, Meng J, Romero P, Yang JY, Chen JW, Vacic V, Obradovic Z, and Uversky VN. 2008. “The Unfoldomics Decade: An Update on Intrinsically Disordered Proteins.” *BMC Genomics* 9 (2): S1.

Dunker AK, and Uversky VN. 2010. “Drugs for ‘Protein Clouds’: Targeting Intrinsically Disordered Transcription Factors.” *Current Opinion in Pharmacology* 10 (6): 782–88.

Dyson HJ, and Wright PE. 2002. “Coupling of Folding and Binding for Unstructured Proteins.” *Current Opinion in Structural Biology* 12 (1): 54–60.

Dyson JH, and Wright PE. 2005. “Intrinsically Unstructured Proteins and Their Functions.” *Nature Reviews Molecular Cell Biology* 6(3):197-208.

Ekman D, Light S, Björklund AK, and Elofsson A. 2006. “What Properties Characterize the Hub Proteins of the Protein-Protein Interaction Network of Saccharomyces Cerevisiae?” *Genome Biology* 7 (6): R45–R45.

Englander SW, Mayne L, Bai Y, and Sosnick TR. 2008. “Hydrogen Exchange: The Modern Legacy of Linderstrøm-Lang.” *Protein Science* 6 (5): 1101–9.

Erkizan HV, Uversky VN, and Toretsky JA. 2010. “Oncogenic Partnerships: EWS-FLI1

Protein Interactions Initiate Key Pathways of Ewing's Sarcoma.” *Clinical Cancer Research* 16 (16): 4077-4083.

Ferron F, Longhi S, Canard B, and Karlin D. 2006. “A Practical Overview of Protein Disorder Prediction Methods.” *Proteins: Structure, Function, and Bioinformatics* 65 (1): 1–14.

Fink AL. 2005. “Natively Unfolded Proteins.” *Current Opinion in Structural Biology* 15 (1): 35–41.

Fisher CK, and Stultz CM. 2011. “Constructing Ensembles for Intrinsically Disordered Proteins.” *Current Opinion in Structural Biology* 21 (3): 426–31.

Fontana A, Polverino de Laureto P, and De Filippis V. 1993. “Molecular Aspects of Proteolysis of Globular Proteins.” *Studies in Organic Chemistry* 47:101–10.

Fontana A, Zambonin M, Polverino de Laureto P, De Filippis V, Clementi A, and Scaramella E. 1997. “Probing the Conformational State of Apomyoglobin by Limited Proteolysis.” *Journal of Molecular Biology* 266 (2): 223–30.

Fonin, A. V., A. L. Darling, I. M. Kuznetsova, K. K. Turoverov and V. N. Uversky (2018). "Intrinsically disordered proteins in crowded milieu: when chaos prevails within the cellular gumbo." *Cell Mol Life Sci* **75**(21): 3907-3929.

Galea CA, Pagala VR, Obenauer JC, Park CG, Slaughter CA, and Kriwacki RW. 2006. “Proteomic Studies of the Intrinsically Unstructured Mammalian Proteome.” *Journal of Proteome Research* 5 (10): 2839–48.

Ghisaidoobe A, and Chung S. 2014. “Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques.” *International Journal of Molecular Sciences* 15 (12): 22518–38.

Ghosh RP, Nikitina T, Horowitz-Scherer RA, Gierasch LM, Uversky VN, Hite K, Hansen JC, and Woodcock CL. 2010. “Unique Physical Properties and Interactions of the Domains of Methylated DNA Binding Protein 2.” *Biochemistry* 49 (20): 4395–4410.

Goh GK, Dunker AK, and Uversky VN. 2008a. “Protein Intrinsic Disorder Toolbox for Comparative Analysis of Viral Proteins.” *BMC Genomics* 9 (Suppl 2): S4–S4.

Goh GK, Dunker AK, and Uversky VN. 2008b. “A Comparative Analysis of Viral Matrix Proteins Using Disorder Predictors.” *Virology Journal* 5: 126.

Gregoret LM, and Sauer RT. 1998. “Tolerance of a Protein Helix to Multiple Alanine and Valine Substitutions.” *Folding & Design* 3 (2): 119–26.

Gustianand M, Liggins JR, Cummins PL, and Gready JE. 2004. “Conformation of Prion Protein Repeat Peptides Probed by FRET Measurements and Molecular Dynamics Simulations.” *Biophysical Journal* 86 (4): 2467–83.

Hammarstro P, and Persson M. 2001. “Protein Compactness Measured by Fluorescence Resonance.” *Journal of Biological Chemistry* 276 (24): 21765–75.

Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, and Iakoucheva LM. 2006. “Intrinsic Disorder Is a Common Feature of Hub Proteins from Four Eukaryotic Interactomes.” *PLOS Computational Biology* 2 (8): e100.

Hegyi H, Buday L, and Tompa P. 2009. “Intrinsic Structural Disorder Confers Cellular Viability on Oncogenic Fusion Proteins.”. *PLoS Computational Biology* 5 (10): e1000552.

Helms MK, Hazlett TL, Mizuguchi H, Hasemann CA, Uyeda K, Jameson DM 1998. "Site-Directed Mutants of Rat Testis Fructose 6-Phosphate, 2-Kinase/Fructose 2,6-Bisphosphatase: Localization of Conformational Alterations Induced by Ligand Binding". *Biochemistry* 37(40):14057–14064.

Hill JJ, and Royer CA. 1997. "Fluorescence approaches to study of protein-nucleic acid complexation". *Methods Enzymol.* 278: 390–416.

Holowka D, Wensel T, and Baird B. 1990. "A nanosecond fluorescence depolarization study on the segmental flexibility of receptor-bound immunoglobulin E". *Biochemistry* 29: 4607-4612.

Homchaudhuri L, De Avila M, Nilsson SB, Bessonov K, Smith GS, Bamm VV, Musse AA, Harauz G, and Boggs JM. 2010. “Secondary Structure and Solvent Accessibility of a Calmodulin-Binding C-Terminal Segment of Membrane-Associated Myelin Basic Protein.” *Biochemistry* 49 (41): 8955–66.

Huang A and Stultz CM. 2008. “The Effect of a ΔK280 Mutation on the Unfolded State of a Microtubule-Binding Repeat in Tau.” *PLOS Computational Biology* 4 (8): e1000155.

Hubbard SJ, Beynon RJ, and Thornton JM. 1998. “Assessment of Conformational Parameters as Predictors of Limited Proteolytic Sites in Native Protein Structures.” *Protein Engineering* 11 (5): 349–59.

Hubbard SJ, Eisenmenger F, and Thornton JM. 1994. “Modeling Studies of the Change in Conformation Required for Cleavage of Limited Proteolytic Sites.” *Protein Science* 3 (5): 757–68.

Iakoucheva LM, Brown CJ, Lawson JD, Obradović Z, and Dunker AK. 2002. “Intrinsic Disorder in Cell-Signaling and Cancer-Associated Proteins.” *Journal of Molecular Biology* 323 (3): 573–84.

Ilawe, N. V., A. E. Raeber, R. Schweitzer-Stenner, S. E. Toal and B. M. Wong (2015). "Assessing backbone solvation effects in the conformational propensities of amino acid residues in unfolded peptides." *Phys Chem Chem Phys* **17**(38): 24917-24924.

Ishima R, and Torchia DA. 2000. “Protein Dynamics from NMR.” *Nature Structural Biology* 7(9):740-3.

Jahn TR, and Radford SE. 2005. “The Yin and Yang of Protein Folding.” *FEBS Journal* 272 (23): 5962–70.

James LC, and Tawfik DS. 2003. “Conformational Diversity and Protein Evolution – a 60-Year-Old Hypothesis Revisited.” *Trends in Biochemical Sciences* 28 (7): 361–68.

James NG, Ross JA, Mason AB, and Jameson DM. 2010. “Excited-State Lifetime Studies of the Three Tryptophan Residues in the N-Lobe of Human Serum Transferrin.” *Protein Science* 19: 99–110.

Jeong H, Mason SP, Barabási AL, and Oltvai ZN. 2001. “Lethality and Centrality in Protein Networks.” *Nature* 411: 41-42.

Jha, A. K., A. Colubri, K. F. Freed and T. R. Sosnick (2005). "Statistical coil model of the unfolded state: resolving the reconciliation problem." *Proc Natl Acad Sci U S A* **102**(37): 13099-13104.

Kamerlin SC, and Warshel A. 2010. “At the Dawn of the 21 St Century : Is Dynamics the Missing Link.” *Proteins: Structure, Function and Bioinformatics* 78 (6): 1339–75.

Karush Fr. 1950. “Heterogeneity of the Binding Sites of Bovine Serum Albumin.” *Journal of the American Chemical Society* 72 (6): 2705–13.

Kast, P., M. Asif-Ullah, N. Jiang and D. Hilvert (1996). "Exploring the active site of chorismate mutase by combinatorial mutagenesis and selection: the importance of electrostatic catalysis." *Proc Natl Acad Sci U S A* **93**(10): 5043-5048.

Khrenova, M., I. Topol, J. Collins and A. Nemukhin (2015). "Estimating orientation factors in the FRET theory of fluorescent proteins: the TagRFP-KFP pair and beyond." *Biophys J* **108**(1): 126-132.

Kiefhaber T, Schmid FX, Willaert K, Engelborghs Y, and Chaffotte A. 1992. "Structure of a Rapidly Formed Intermediate in RNase Tl Folding." *Protein Science* 1(9):1162-72.

Kim TD., Ryu HY, Cho HI, Yang CH, and Kim J. 2000. “Thermal Behavior of Proteins: Heat-Resistant Proteins and Their Heat-Induced Secondary Structural Changes.” *Biochemistry* 39 (48): 14839–46.

Koshland DE. 1958. “Application of a Theory of Enzyme Specificity to Protein Synthesis.” *Proceedings of the National Academy of Sciences* 44 (2): 98–104.

Koshland DE. 1995. “The Key–Lock Theory and the Induced Fit Theory.” *Angewandte Chemie International Edition in English* 33 (23–24): 2375–78.

Kriwacki RW, Wu J, Tennant L, Wright PE, and Siuzdak G. 1997. “Probing Protein Structure Using Biochemical and Biophysical Methods: Proteolysis, Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry, High-Performance Liquid Chromatography and Size-Exclusion Chromatography of P21Waf1/Cip1/Sdi1.” *Journal of Chromatography A* 777 (1): 23–30.

Kuwajima K. 1977. “A Folding Model of α-Lactalbumin Deduced from the Three-State Denaturation Mechanism.” *Journal of Molecular Biology* 114 (2): 241–58.

Kuwajima K. 1989. “The Molten Globule State as a Clue for Understanding the Folding and Cooperativity of Globular-Protein Structure.” *Proteins: Structure, Function, and Bioinformatics* 6 (2): 87–103.

Lakowicz J. 2006. *Principles of Fluorescent Spectroscopy, 3rd Edn*. Vol. 1.

Lakowicz JR, Gryczynski I, Laczko G, Wiczk W, and Johnson ML. 1994. “Distribution of Distances between the Tryptophan and the N-terminal Residue of Melittin in Its Complex with Calmodulin, Troponin C, and Phospholipids.” *Protein Science* 3 (4): 628–37.

Langridge, T. D., M. J. Tarver and S. T. Whitten (2014). "Temperature effects on the hydrodynamic radius of the intrinsically disordered N-terminal region of the p53 protein." *Proteins* **82**(4): 668-678.

Lee AY, Karplus PA, Ganem B, and Clardy J. 1995. "Atomic Structure of the Buried Catalytic Pocket of Escherichia Coli Chorismate Mutase." *Journal of the American Chemical Society* 117 (12): 3627–28.

Lee, A. Y., J. D. Stewart, J. Clardy and B. Ganem (1995). "New insight into the catalytic mechanism of chorismate mutases from structural studies." *Chem Biol* **2**(4): 195-203.

Li YH, and Jing GZ. 2000. "Double Point Mutant F34W/W140F of Staphylococcal Nuclease Is in a Molten Globule State but Highly Competent to Fold into a Functional Conformation." *Journal of Biochemistry* 128(5):739-44.

Ma L, Yang F, and Zheng J. 2014. "Application of Fluorescence Resonance Energy Transfer in Protein Studies." *Journal of Molecular Structure* 1077: 87–100.

MacBeath G, Kast P, and Hilvert D. 1998. "A Small, Thermostable, and Monofunctional Chorismate Mutase from the Archeon Methanococcus Jannaschii." *Biochemistry* 37 (28): 10062–73.

Macbeath G, Kast P, and Hilvert D. 1998. "Redesigning Enzyme Topology by Directed Evolution." *Science* 279 (5358): 1958–61.

Mao AH, Crick SL, Vitalis A, Chicoine CL, and Pappu RV. 2010. "Net Charge per Residue Modulates Conformational Ensembles of Intrinsically Disordered Proteins." *Proceedings of the National Academy of Sciences* 107 (18): 8183 LP-8188.

Markus G. 1965. "Protein Substrate Conformation and Proteolysis." *Proceedings of the National Academy of Sciences* 54 (1): 253 LP-258.

Meer BW, Coker III G & Simon Chen SY (1994) 'Resonance energy transfer: theory and data', VCH, New York.

Midic U, Oldfield CJ, Dunker AK, Obradovic Z, and Uversky VN. 2009. "Protein Disorder in the Human Diseasome: Unfoldomics of Human Genetic Diseases." *BMC Genomics* 10 (Suppl 1): S12–S12.

Mittag T, Kay LE, and Forman-Kay JD. 2010. "Protein Dynamics and Conformational Disorder in Molecular Recognition." *Journal of Molecular Recognition* 23 (2): 105–16.

Mittag T, Marsh J, Grishaev A, Orlicky S, Lin H, Sicheri F, Tyers M, and Forman-Kay JD. 2010. "Structure/Function Implications in a Dynamic Complex of the Intrinsically Disordered Sic1 with the Cdc4 Subunit of an SCF Ubiquitin Ligase." *Structure* 18 (4): 494–506.

Mittag T, Orlicky S, Choy WY, Tang X, Lin H, Sicheri F, Kay LE, Tyers M, and Forman-Kay JD. 2008. "Dynamic Equilibrium Engagement of a Polyvalent Ligand with a Single-Site Receptor." *Proceedings of the National Academy of Sciences* 105 (46): 17772 LP-17777.

Miyawaki A. 2011. "Development of Probes for Cellular Functions Using Fluorescent Proteins and Fluorescence Resonance Energy Transfer." *Annual Review of Biochemistry*

80 (1): 357–73.

Möglich A, Joder K, and Kiefhaber T. 2006. “End-to-End Distance Distributions and Intrachain Diffusion Constants in Unfolded Polypeptide Chains Indicate Intramolecular Hydrogen Bond Formation.” *Proceedings of the National Academy of Sciences* 103 (33): 12394 LP-12399.

Mohan A, Sullivan Jr J, Radivojac P, Keith A, and Uversky VN. 2008. “Intrinsic Disorder in Pathogenic and Non-Pathogenic Microbes : Discovering and Analyzing the Unfoldomes of Early-Branching Eukaryotes.” *Molecular BioSystems* 4 (4): 328–40.

Moosa, M. M., A. C. Ferreon and A. A. Deniz (2015). "Forced folding of a disordered protein accesses an alternative folding landscape." *Chemphyschem* **16**(1): 90-94.

Muchmore SW, Sattler M, Liang H, Meadows RP, Harlan JE, Yoon HS, Nettesheim D, Chang BS, Thompson CB, Wong SL, Ng SL, Fesik SW. 1996. “X-Ray and NMR Structure of Human Bcl-XL, an Inhibitor of Programmed Cell Death.” *Nature* 381(6580):335-41.

Mukhopadhyay S, Rajaraman K, Lemke EA, Lindquist S, and Deniz AA. 2007. “A Natively Unfolded Yeast Prion Monomer Adopts an Ensemble of Collapsed and Rapidly Fluctuating Structures.” *Proceedings of the National Academy of Sciences* 104 (8): 2649 LP-2654.

Nagano N, Ota M, and Nishikawa KY. 1999. “Strong Hydrophobic Nature of Cysteine Residues in Proteins” *FEBS Letters* 458(1):69-71.

Narayanan RL, Dürr UHN, Bibow S, Biernat J, Mandelkow E, and Zweckstetter M. 2010. “Automatic Assignment of the Intrinsically Disordered Protein Tau with 441-Residues.” *Journal of the American Chemical Society* 132 (34): 11906–7.

Nash P, Tang X, Orlicky S, Chen Q, Gertler FB, Mendenhall MD, Sicheri F, Pawson T, and Tyers M. 2001. “Multisite Phosphorylation of a CDK Inhibitor Sets a Threshold for the Onset of DNA Replication.” *Nature* 414 (November): 514.

Nelson M and McClelland MB.1992. "Use of DNA methyltransferase/endonuclease enzyme combinations for megabase mapping of chromosomes" *Methods Enzymology* 216:279-303.

Ohgushi M, and Wada A. 1983. “‘Molten-Globule State’: A Compact Form of Globular Proteins with Mobile Side-Chains.” *FEBS Letters* 164 (1): 21–24.

Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, and Dunker AK. 2005. “Comparing and Combining Predictors of Mostly Disordered Proteins.” *Biochemistry* 44 (6): 1989–2000.

Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, and Dunker AK. 2005. “Coupled Folding and Binding with α-Helix-Forming Molecular Recognition Elements.” *Biochemistry* 44 (37): 12454–70.

Papworth C, Bauer JC, and Braman J. 1996. "Site-Directed Mutagenesis in One day with >80% Efficiency," *Strategies*.

Pappu, R. V., R. Srinivasan and G. D. Rose (2000). "The Flory isolated-pair hypothesis is not valid for polypeptide chains: implications for protein folding." *Proc Natl Acad Sci U S A* **97**(23): 12565-12570.

Patil A, and Nakamura H. 2006. “Disordered Domains and High Surface Charge Confer Hubs with the Ability to Interact with Multiple Proteins in Interaction Networks.” *FEBS Letters* 580 (8): 2041–45.

Peracchi, A. (2001). "Enzyme catalysis: removing chemically 'essential' residues by site-directed mutagenesis." *Trends Biochem Sci* **26**(8): 497-503.

Pervushin K, Vamvaca K, Vögeli B, and Hilvert D. 2007. “Structure and Dynamics of a Molten Globular Enzyme.” *Nature Structural and Molecular Biology* 14 (12): 1202–6.
Plaxco KW, Groβ M. 1997. “The Importance of Being Unfolded.” *Nature* 386: 657–659

Pontius BW. 1993. “Close Encounters: Why Unstructured, Polymeric Domains Can Increase Rates of Specific Macromolecular Association.” *Trends in Biochemical Sciences* 18 (5): 181–86.

Popelka H, Uversky VN, and Klionsky DJ. 2014. “Identification of Atg3 as an Intrinsically Disordered Polypeptide Yields Insights into the Molecular Dynamics of Autophagy-Related Proteins in Yeast.” *Autophagy* 10 (6): 1093–1104.

Prasad S, Mandal I, Singh S, Paul A, Mandal B, Venkatramani R, Swaminathan R. 2017. "Chemical Science Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein", *Chemical Science* 8: 5416–5433.

Ptitsyn OB, Pain RH, Semisotnov GV, Zerovnik E, and Razgulyaev OI. 1990. “Evidence for a Molten Globule State as a General Intermediate in Protein Folding.” *FEBS Letters* 262 (1): 20–24.

Radford SE. 2000. “Protein Folding: Progress Made and Promises Ahead.” *Trends in Biochemical Sciences* 25 (12): 611–18.

Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, and Dunker AK. 2007. “Intrinsic Disorder and Functional Proteomics.” *Biophysical Journal* 92 (5): 1439–56.

Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, Brown CJ, Lawson JD, and Dunker AK. 2004. “Protein Flexibility and Intrinsic Disorder.” *Protein Science* 13 (1): 71–80.

Roland R, Hornemann S, Wider G, Billeter M, Glockshuber R, and Wüthrich K. 1996. “NMR Structure of the Mouse Prion Protein Domain PrP(121–231).” *Nature* 382(6587):180-2

Romero P, Obradovic Z, Kissinger CR, Villafranca JE, Garner E, Guilliot S, and Dunker AK. 1998. “Thousands of Proteins Likely to Have Long Disordered Regions.” *Pacific Symposium on Biocomputing* 437–48.

Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, and Dunker AK. 2001. “Sequence Complexity of Disordered Protein.” *Proteins: Structure, Function, and Bioinformatics* 42 (1): 38–48.

Rusinova E, Tretyachenko-Ladokhina V, Vele OE, Senear DF, and Alexander Ross JB .2002. "Alexa and Oregon Green dyes as fluorescence anisotropy probes for measuring protein–protein and protein–nucleic acid interactions". *Anal. Biochem* 308: 18-25.

Sambrook J, Fritsch EF, and Maniatis T (1989): Molecular cloning: a laboratory manual. *Cold spring harbor laboratory press.*

Schweitzer-Stenner, R. (2012). "Conformational propensities and residual structures in unfolded peptides and proteins." *Mol Biosyst* **8**(1): 122-133.

Schweitzer-Stenner, R. and S. E. Toal (2016). "Construction and comparison of the statistical coil states of unfolded and intrinsically disordered proteins from nearest-neighbor corrected conformational propensities of short peptides." *Mol Biosyst* **12**(11): 3294-3306.

Schweitzer-Stenner, R. and S. E. Toal (2018). "Anticooperative Nearest-Neighbor Interactions between Residues in Unfolded Peptides and Proteins." *Biophys J* **114**(5): 1046-1057.

Schweers O, Schönbrunn-Hanebeck E, Marx A, and Mandelkow E. 1994. "Structural Studies of Tau Protein and Alzheimer Paired Helical Filaments Show No Evidence for Beta-Structure." *The Journal of Biological Chemistry* 269 (39): 24290–97.

Semisotnov GV, Rodionova NA, Razgulyaev OI, Uversky VN, Gripas' AF, and Gilmanshin RI. 1991. "Study of the 'Molten Globule' Intermediate State in Protein Folding by a Hydrophobic Fluorescent Probe." *Biopolymers* 31 (1): 119–28.

She M, Xing J, Dong WJ, Umeda PK, and Cheung HC. 1998. "Calcium Binding to the Regulatory Domain of Skeletal Muscle Troponin C Induces a Highly Constrained Open Conformation." *Journal of Molecular Biology* 281 (3): 445–52.

Sieber T, Scholz R, Spoerner M, Schumann F, Kalbitzer HR, and Dobner T. 2011. "Intrinsic Disorder in the Common N-Terminus of Human Adenovirus 5 E1B-55K and Its Related E1BN Proteins Indicated by Studies on E1B-93R." *Virology* 418 (2): 133–43.

Sigalov AB. 2010. "Protein Intrinsic Disorder and Oligomericity in Cell Signaling." *Molecular BioSystems* 6 (3): 451–61.

Singh GP, and Dash D. 2007. "Intrinsic Disorder in Yeast Transcriptional Regulatory Network." *Proteins: Structure, Function, and Bioinformatics* 68 (3): 602–5.

Singh GP, Ganapathi M, and Dash D. 2007. "Role of Intrinsic Disorder in Transient Interactions of Hub Proteins." *Proteins: Structure, Function, and Bioinformatics* 66 (4): 761–65.

Small EW, and Isenberg I. 1977. "Hydrodynamic properties of a rigid molecule: Rotational and linear diffusion and fluorescence anisotropy." *Biopolymers* 16: 1907-1928.

Spolar RS, and Record MT. 1994. "Coupling of Local Folding to Site-Specific Binding of Proteins to DNA." *Science* 263 (5148): 777 LP-784.

Stryer L. 1965. "The interaction of a naphthalene dye with apomyoglobin and apohemoglobin: A fluorescent probe of non-polar binding sites". *Journal of Molecular Biology* 13(2): 482–495.

Stryer L. 1967. “Fluorescence Energy Transfer as a Spectroscopic Ruler.” *Annual Review of Biochemistry* 58 (2): 719–26.

Swaminathan R, Periasamy N, Udgaonkar JB, and Krishnamoorthy G.1994. "Molten Globule-like Conformation of Barstar: A Study by Fluorescence Dynamics." *J. Phys. Chem.* 98: 9270-9278.

Toal, S. and R. Schweitzer-Stenner (2014). "Local order in the unfolded state: conformational biases and nearest neighbor interactions." Biomolecules **4**(3): 725-773.

Tompa P. 2002. “Intrinsically Unstructured Proteins.” *Trends in Biochemical Sciences* 27 (10): 527–33.

Toscano, M. D., K. J. Woycechowsky and D. Hilvert (2007). "Minimalist active-site redesign: teaching old enzymes new tricks." *Angew Chem Int Ed Engl* **46**(18): 3212-3236.

Tsvetkov P, Myers N, Moscovitz O, Sharon M, Prilusky J, and Shaul Y. 2012. “Thermo-Resistant Intrinsically Disordered Proteins Are Efficient 20S Proteasome Substrates.” *Molecular BioSystems* 8 (1): 368–73.

Turoverov KK, Kuznetsova IM, and Uversky VN. 2010. “The Protein Kingdom Extended: Ordered and Intrinsically Disordered Proteins, Their Folding, Supramolecular Complex Formation, and Aggregation.” *Progress in Biophysics and Molecular Biology* 102 (2): 73–84.

Uversky VN, and Dunker AK. 2010. "Understanding Protein Non-Folding". *Biochimica et Biophysica Acta* 1804: 1231-1264.

Uversky VN, Kutyshenko V, NYu P, Rogov V, Vassilenko K, and Gudkov A. 1996. "Circularly Permuted Dihydrofolate Reductase Possesses All the Properties of the Molten Globule State, but Can Resume Functional Tertiary Structure by Interaction with Its Ligands". *Protein Science : A Publication of the Protein Society*. Vol. 5.

Uversky VN. 2011. “Intrinsically Disordered Proteins from A to Z.” *International Journal of Biochemistry and Cell Biology* 43 (8): 1090–1103.

Uversky VN. 2002a. “Natively Unfolded Proteins: A Point Where Biology Waits for Physics.” *Protein Science* 11 (4): 739–56.

Uversky VN. 2002b. “What Does It Mean to Be Natively Unfolded?” *European Journal of Biochemistry* 269 (1): 2–12.

Uversky VN. 2013. “Unusual Biophysics of Intrinsically Disordered Proteins.” *Biochimica et Biophysica Acta - Proteins and Proteomics* 1834 (5): 932–51.

Uversky VN, Gillespie JR, and Fink AL. 2000. “Why Are ‘Natively Unfolded’ Proteins Unstructured under Physiologic Conditions?” *Proteins: Structure, Function, and Bioinformatics* 41 (3): 415–27.

Uversky VN, Oldfield CJ, and Dunker AK. 2008. “Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept.” *Annual Review of Biophysics* 37 (1): 215–46.

Uversky VN, Roman A, Oldfield CJ, and Dunker AK. 2006. “Protein Intrinsic Disorder and Human Papillomaviruses: Increased Amount of Disorder in E6 and E7 Oncoproteins from High Risk HPVs.” *Journal of Proteome Research* 5(8): 1829–42.

Uversky VN, Oldfield CJ, and Dunker AK. 2005. "Showing Your ID: Intrinsic Disorder as an ID for Recognition, Regulation and Cell Signaling*". Journal of Molecular Recognition* 18(5): 343-384

Uversky, V. N. (2019). "Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics." *Frontiers in Physics* **7**(10).

Uversky, V. N. (2019). "Protein intrinsic disorder and structure-function continuum." *Prog Mol Biol Transl Sci* **166**: 1-17.

Uversky, V. N., C. Santambrogio, S. Brocca and R. Grandori (2012). "Length-dependent compaction of intrinsically disordered proteins." *FEBS Lett* **586**(1): 70-73.

Vacic V, Uversky VN, Dunker AK, and Lonardi S. 2007. “Composition Profiler: A Tool for Discovery and Visualization of Amino Acid Composition Differences.” *BMC Bioinformatics* 8 (1): 211.

Vamvaca K, Vogeli B, Kast P, Pervushin K, and Hilvert D. 2004. “An Enzymatic Molten Globule: Efficient Coupling of Folding and Catalysis.” *Proceedings of the National Academy of Sciences* 101 (35): 12860–64.

Vamvaca K, Jelesarov I, and Hilvert D. 2008. “Kinetics and Thermodynamics of Ligand Binding to a Molten Globular Enzyme and Its Native Counterpart.” *Journal of Molecular Biology* 382 (4): 971–77.

Vamvaca K, Vögeli B, Kast P, Pervushin K, and Hilvert D. 2004. “An Enzymatic Molten Globule: Efficient Coupling of Folding and Catalysis.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (35): 12860 LP-12864.

van der Lee, R., M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright and M. M. Babu (2014). "Classification of intrinsically disordered regions and proteins." *Chem Rev* **114**(13): 6589-6631.

Vendruscolo M. 2010. “Enzymatic Activity in Disordered States of Proteins.” *Current Opinion in Chemical Biology* 14 (5): 671–75.

Vidalain PO, and Tangy F. 2010. “Virus-Host Protein Interactions in RNA Viruses.” *Microbes and Infection* 12 (14): 1134–43.

Vitalis A, Wang X, and Pappu RV. 2007. “Quantitative Characterization of Intrinsic Disorder in Polyglutamine: Insights from Analysis Based on Polymer Theories.” *Biophysical Journal* 93 (6): 1923–37.

Vivian JT, and Callis PR. 2001. “Mechanisms of Tryptophan Fluorescence Shifts in Proteins.” *Biophysical Journal* 80 (5): 2093–2109.

Walters RH, and Murphy RM. 2009. “Examining Polyglutamine Peptide Length: A Connection between Collapsed Conformations and Increased Aggregation.” *Journal of Molecular Biology* 393 (4): 978–92.

Wang X, Vitalis A, Wyczalkowski MA, and Pappu RV. 2006. “Characterizing the Conformational Ensemble of Monomeric Polyglutamine.” *Proteins: Structure, Function, and Bioinformatics* 63 (2): 297–311.

Warshel A, Sharma P, Kato M, Xiang Y, Liu H, and Olsson M. 2006. "Electrostatic Basis for Enzyme Catalysis." *Chemical Reviews* 106(8):3210-35.

Weber G. 1960. "Fluorescence-polarization spectrum and electronic-energy transfer in tyrosine, tryptophan and related compounds". *Biochemical Journal* 75(2): 335 LP-345.

Weber G, and Shinitzky M. 1970. "Failure of Energy Transfer between Identical Aromatic Molecules on Excitation at the Long Wave Edge of the Absorption Spectrum". *Proceedings of the National Academy of Sciences of the United States of America* 65(4): 823–830.

Weinreb PH, Weiguo Z, Poon AW, Conway KA, and Lansbury PT. 1996. “NACP, A Protein Implicated in Alzheimer’s Disease and Learning, Is Natively Unfolded.” *Biochemistry* 35 (43): 13709–15.

Williams DH, Stephens E, and Zhou M. 2003. “Ligand Binding Energy and Catalytic Efficiency from Improved Packing within Receptors and Enzymes.” *Journal of Molecular Biology* 329 (2): 389–99.

Wiederschain GY. 2011. "The Molecular Probes handbook. A guide to fluorescent probes and labeling technologies." *Springer.*

Williams RM, Obradovi Z, Mathura V, Braun W, Garner EC, Young J, Takayama S, Brown CJ, and Dunker AK. 2001. “The Protein Non-Folding Problem: Amino Acid Determinants of Intrinsic Order and Disorder.” *Pacific Symposium on Biocomputing* 100: 89–100.

Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, and Kern D. 2004. “Linkage between Dynamics and Catalysis in a Thermophilic-Mesophilic Enzyme Pair.” *Nature Structural & Molecular Biology* 11(10):945-9.

Wolfenden R, and Snider MJ. 2002. "The Depth of Chemical Time and the Power of Enzymes as Catalysts." *Accounts of Chemical Research* 34(12):938-45.

Woycechowsky, K. J., A. Choutko, K. Vamvaca and D. Hilvert (2008). "Relative tolerance of an enzymatic molten globule and its thermostable counterpart to point mutation." *Biochemistry* **47**(51): 13489-13496

Wright PE, and Dyson HJ. 1999. “Intrinsically Unstructured Proteins: Re-Assessing the Protein Structure-Function Paradigm.” *Journal of Molecular Biology* 293 (2): 321–31.

Wu PG, and Brand L. 1994. "Resonance energy transfer: methods and applications." *Anal. Biochem.* 218:1-13.

Xue B, Williams RW, Oldfield CJ, Goh GK, Dunker AK, and Uversky VN. 2010. “Viral Disorder or Disordered Viruses: Do Viral Proteins Possess Unique Features?” *Protein & Peptide Letters* 17(8):932-51.

Yamada J, Phillips JL, Patel S, Goldfien G, Calestagne-Morelli A, Huang H, Reza R, et al. 2010. “A Bimodal Distribution of Two Distinct Categories of Intrinsically Disordered Structures with Separate Functions in FG Nucleoporins.” *Molecular & Cellular Proteomics* 9 (10): 2205 LP-2224.

Yang C, van der Woerd MJ, Muthurajan UM, Hansen JC, and Luger K. 2011. “Biophysical Analysis and Small-Angle X-Ray Scattering-Derived Structures of MeCP2–nucleosome Complexes.” *Nucleic Acids Research* 39 (10): 4122–35.

Yguerabide J, Epstein HF, and Stryer L (1970): Segmental flexibility in an antibody molecule. *J. Mol. Biol.* **51**, 573-590.

Zaman, M. H., M. Y. Shen, R. S. Berry, K. F. Freed and T. R. Sosnick (2003). "Investigations into sequence and conformational dependence of backbone entropy, inter-basin dynamics and the Flory isolated-pair hypothesis for peptides." *J Mol Biol* **331**(3): 693-711.

Zhou HX. 2012. “Intrinsic Disorder: Signaling via Highly Specific but Short-Lived Association.” *Trends in Biochemical Sciences* 37 (2): 43–48.

List of conferences attended and publications

Biswas S, Swaminathan R., Investigation of the Structure and Dynamics of Intrinsically Disordered Enzymes using Fluorescence Based Techniques, FCS (IISc), Bangalore, 2013.

Attended a Workshop on “Statistical Data Analysis” at Indian Statistical Institute (ISI), Kolkata, September 2014

Biswas S, Swaminathan R., Intrinsically Disordered Monomeric Chorismate Mutase: Purification, Mutagenesis and Structural Investigations, Journal of Proteins and Proteomics 5(5): JPP138, 2014.

Biswas S, Rao AG, Bhattacharya S, Swaminathan R., Investigating the Structural and Functional Dynamics of Intrinsically Disordered Enzyme by Fluorescence Spectroscopy and MD Simulations, GRC IDP, Switzerland, 2016

Biswas S, Rao AG, Bhattacharya S, Swaminathan R., Probing the Local dynamics of an Intrinsically Disordered Molten Globule Enzyme by Fluorescence Spectroscopy and MD simulations, OWL (TIFR), Mumbai, 2016.

Biswas S, Swaminathan R., Insights into the Functional Dynamics of an Intrinsically Disordered Enzyme using Fluorescence Based Tools, Research Conclave (IIT Guwahati), 2019

Chandra N, Biswas S, Rout J, Basu G, Tripathy U. Stability of β-turn in LaR2C-N7 peptide for its translation-inhibitory activity against hepatitis C viral infection: A molecular dynamics study. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy (2019), 211:26-33.

Biswas S, Gavra VB, Das AK, and Tripathy U. Biophotonics in disease diagnosis and therapy. Biomedical Engineering and its application in healthcare. Springer Nature 2019. [book chapter]

Biswas S, Gangadhar P, Srihari P, Swaminathan R. Time-Resolved Fluorescence Coupled with Site-Directed Mutagenesis reveals Functional Dynamics of an Intrinsically Disordered Enzyme, Monomeric Chorismate Mutase (Manuscript under preparation)