Spectroscopic, structural and functional characterization of Intrinsically Disordered Protein DHN1 from *Zea mays*

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

By

Dileep Ahari

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

August 2022

TH-3004_146106002

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 and Prof. Latha Rangan.

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
August, 2022

Dileep Ahari
Roll no. 146106002

TH-3004_146106002

INDIAN INSTITUTE OF TECHNOLOGY

GUWAHATI, Assam, India

Department of Biosciences and Bioengineering

CERTIFICATE

It is certified that the work described in this thesis, entitled "***Spectroscopic, structural and functional characterization of Intrinsically Disordered Protein DHN1 from Zea mays***" done by **Dileep Ahari** 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
August, 2022

Prof. Rajaram Swaminathan & Prof. Latha Rangan
Department of BSBE
IIT Guwahati

TH-3004_146106002

Dedicated to my family

and friends

TH-3004_146106002

ACKNOWLEDGEMENTS

First and foremost, I want to express my deepest respect and sincere gratitude to my both supervisor Prof. R. Swaminathan and Prof. Latha Rangan, for giving me the opportunity to undertake a Ph.D. in their group. Their continuous support, patience, enthusiasm, immense scientific ideas and motivation helped me in all the time of research and to explore the domain of workassembled in this thesis. Especially Prof. Rajaram Swaminathan, who has been a great mentor, from whom I have learned a lot of things. He encouraged me not only to grow as a researcher but also as a person. I am indebted to him, for his faith in me and providing the right direction whenever I needed it the most.

Besides my advisor, I would also like to extend my sincere thanks to the Doctoral Committee members, Prof. K. Pakshirajan, Dr. Manish Kumar and Dr. Sunanda Chatterjee for their encouragement, timely evaluation of my Ph.D. work and insightful suggestions, which help me a lot in the betterment of my thesis.

I thank Prof. Timothy J. Close at University of California, Riverside for kindly providing us the DHN1 plasmid construct.

Thanks to the BSBE department staff members, for their co-operative nature. My sincere thanks to Central Instruments Facility, for service and support to Mass spectrometryfacilities.

Thanks to IIT Guwahati and its administration for all the facilities that were made available to me and the Ministry of Human Resource Development (MHRD, India for the financial support. I would like to thank IIT Guwahati to support me financially in the 6th year of my Ph.D.

Next, I wish to express my sincere thanks to my lab seniors Dr. Ziauddin Ansari, Dr. Amrendra Kumar, Dr. Saumya Prasad, Dr. Shrutidhara Biswas and Dr. Anuma,. for their enormous help and suggestion during my Ph.D. For making the lab a great place to work, I express my deep sense of gratitude to all my extraordinary juniors Anurag, Ekramul, Sanjana, Alka, Heeramoni, Simangka, Nuzelu, Himanshi, Vinay, Nayan, Aditya, Garrema, Subhajit, Abheek. Harikrishna, Nitish and Kashish Sahil.

TH-3004_146106002

I am especially thankful to Ekramul and Simangka for their help during the thesis writing.

I would also like to express my sincere thanks to all my friends, Bhuvan, Suraj, Bapi, Abshar, Amit and Minati, for their constant support and timely help. I thank them for making my stay at IIT Guwahati a memorable one.

My Ph.D. endeavor could not have been completed without the endless love, unending support, tolerance and blessings from my family. My deepest gratitude to my parents, Late Smt. Anita Devi and Shri. Shambhulal Ahari whose enormous sacrifices, unconditional love in every stage of my life motivated me overcomes all the challenges. Especially father, whose firm believes in me and encouragement has given me all the strength to break through all the difficulties in life.I am also grateful to my brother and sisters, Vijay, Radha and Khusbhoo for their unending love and emotional support in all the times. I owe every bit of my achievement to my family.

There were many other people who helped me both personally and professionally. Although it is not possible to pen down each of their names. I would like to thank each and every one of them for helping me in some or the other way during the different stages of my Ph.D.

Last but not the least; I am thankful to almighty God for his continuous blessing during my research career to achieve these remarkable steps.

Dileep Ahari *August, 2022*

TH-3004_146106002

LIST OF ABBREVIATIONS

NATA	N-acetyl tryptophan amide
IPTG	Isopropyl β-D-1-thiogalactopyranoside
TFE	2,2,2-Trifluoroethanol
IAEDANS	(1,5-IAEDANS,5-((((2-Iodoacetyl)amino)ethyl)amino)Naphthalene-1- Sulfonic acid)
FRET	Förster Resonance Energy Transfer
DTT	1,4-Dithiotreotol
BSA	Bovine Serum Albumin
TCEP	Tris(2-carboxyethyl)phosphine
MALDI	Matrix Assisted Laser Desorption/Ionization
CD	Circular Dichroism
SDS	Sodium Dodecyl Sulphate
FWHM	Full-width half-maximum
IDP	Intrinsically Disordered Protein
IDR	Intrinsically Disordered Region
MEM	Maximum entropy method
DMF	Dimethylformamide
dsDNA	double-Stranded deoxyribonucleic acid
PCR	Polymerase chain reaction
HEWL	Hen Egg White Lysozyme
LDH	Lactate Dehydrogenase

TH-3004_146106002

HuSA	Human Serum Albumin
ProCharTS	Protein Charge Transfer Spectra
Ni-NTA	Nickel-Nitrilotriacetic acid
UV-Vis	Ultraviolet-Visible
Dansyl	2-dimethyl aminonaphthalene 6-sulfonyl
PAGE	Polyacrylamide Gel Electrophoresis

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी • Indian Institute of Technology Guwahati •

TH-3004_146106002

Thesis Abstract

Among all the macromolecules, protein has been shown to have a critical and diverse role in maintaining cellular life. Intrinsically disordered proteins (IDPs) emerged as an essential class of proteins in the past two decades due to their functional relevance without having a proper 3-dimensional structure. They contrast to the older structure-function paradigm, where the specific 3D structure of proteins is required for its proper functioning. This protein class has been shown to have a crucial role in various cellular processes like signaling, regulation and control.

Here, we investigated the DHN1 (Dehydration protein) protein from *Zea mays,* a member of group 2 late embryogenesis abundant protein family. Several studies suggested that this protein upregulates during abiotic stress like drought, cold and salinity. *In vitro* experiments, evidence from biochemical assays and localization experiments suggests multiple roles for dehydrins such as membrane protection, cryoprotection of enzyme and protection from reactive oxygen species.

This study is focused on studying DHN1 structural dynamics and functional role using the novel spectroscopic tool Protein Charge Transfer Spectra (ProCharTS) and other spectroscopic techniques like UV-visible and Fluorescence spectroscopy, Circular Dichroism (CD) and Förster resonance energy transfer (FRET). To investigate the structural change and functional role of DHN1, two mutants: DHN1 CW1 (Trp^{122}—Cys^{62}) and DHN1 W3 (Trp^{3}) were generated using site-directed mutagenesis. DHN1 and its mutants were expressed as recombinant protein in *E. coli* and purified.

The characterization of the intrinsic disorder of DHN1 and its mutant proteins in the native state was done using the sequence-based predictor experimentally. We observed anomalous mobility in reducing SDS PAGE and heat resistance of DHN1 and its mutants. The steady state fluorescence emission, and fluorescence anisotropy revealed the solvent-exposed tryptophan in DHN1 CW1 and DHN1 W3 mutant. The secondary structure content of DHN1 and its mutant proteins was confirmed by circular dichroism (CD). Structural analysis revealed that substitution of tryptophan and cysteine in DHN1 protein does not alter the structural features of DHN1.

TH-3004_146106002

Next, we exploited the richness of charged amino acids in DHN1 protein to investigate the structural dynamics of DHN1 protein using the ProCharTS. ProCharTS originates due to photoinduced charge transfer from: polypeptide backbone to NH_3^+ in the lysine; COO^- in glutamate to polypeptide backbone; and COO^- in glutamate to NH_3^+ in lysine. Conformational changes induced in the DHN1 protein by altering the temperature and pH of the aqueous medium were monitored using the ProCharTS. We observed the substantial change in the ProCharTS with changing temperature (25 −85 °C) and pH (3-11). Further, the presence of 250 mM NaCl or KCl in the medium also altered the ProCharTS spectrum.

Subsequently, we looked into the photochemical feature of intrinsic luminescence arising from the charge transfer state of ProCharTS in monomeric DHN1. This novel ProCharTS are emissive in nature throughout the UV-Visible region. However, the luminescent yield at the longer wavelength regions is found to be much lower in comparison to the shorter wavelengths. The ProCharTS luminescence from the charge transfer states in monomeric DHN1 protein has lower quantum yield (~0.027 at λ_{ex} 355 nm) and the low mean luminescence lifetime (~2.8 ns at 340 nm). However, the ProCharTS luminescence shows the huge Stokes shift($> 6942\ to\ > 3159$ cm^{-1} with increase in λ_{ex} from 310 to 410 nm). Due to the prominent ProCharTS luminescence throughout the UV-Visible region, the potential influence of ProCharTS on the intrinsic fluorophore like Trp in proteins is also investigated. The presence of high charged residues in DHN1 protein quenched the fluorescence of NATA and also changed the nature of fluorophore intensity decay of NATA from single exponential to two exponentials. However, fluorescence lifetime distribution of Trp in DHN1 was not influenced by presence of charged amino acids. This suggests that ProCharTS can sometimes alter the fluorescence emission of Trp in protein rich in charged amino acids.

Further, we studied the DHN1 protein in the presence of 2,2,2 trifluoroethanol (TFE), a well-known helix-inducing solvent that is reported to mimic the membrane environment; at high temperatures by using circular dichroism. A disorder-to-order transitions were observed in DHN1 and its mutant proteins with increase in TFE and temperature.

An investigation of the DHN1 protein also performed in the presence of SDS in monomer and at a micellar concentration to understand the role of conserved segments and N-terminal in DHN1 protein folding. DHN1 and its mutant DHN1 CW1 and DHN1 W3 in the presence of SDS were studied by using steady state fluorescence, fluorescence anisotropy, circular dichroism and time-resolved fluorescence measurement. To analyze the overall structural

TH-3004_146106002

changes in DHN1 protein at different concentration of SDS, Förster Resonance Energy Transfer (FRET) was carried out. Our spectroscopic and FRET data suggest the folding of two K-segment in DHN1 closer to C-terminal in presence of SDS. Further, an observation was also made on the role of the N-terminal in the folding of DHN1 protein in the presence of SDS. Further, the cryoprotection and heat protection activity of DHN1 and its mutants were measured. However, the functional properties of the DHN1 mutant were slightly altered, and the cryoprotection activity of the DHN1 mutants is slightly less compared to the wildtype DHN1 protein. At the same time, the heat protection activity increased in both the mutants as compared to DHN1, where both mutants protect the LDH enzyme denaturation or degradation when incubated to a higher than physiological temperature. DHN1 protein showed weak electrostatic binding/interaction to LDH enzyme, which could play a potential role in the functional mechanism of DHN1 in stressed conditions.

Taken together, the current thesis work revealed the structural dynamics and the disorder-to-order transitions of DHN1 protein. ProCharTS can serve as a label-free tool to study and detect structural transitions in the IDPs which are rich in charged amino acid residues and devoid of aromatic chromophores. This novel ProCharTS is emissive in nature and persist all over the UV-visible region. Due to its presence throughout the UV-Visible region, it can affect the fluorescence of the other intrinsic fluorophore present in the protein. Further, the DHN1 protein shows the folding in the presence of SDS at monomeric concentration, where the two K-segment come close to the C-terminal of DHN1 and initiate the folding later, followed by the N-terminal of DHN1 protein. Finally, the functional studies of DHN1 protein indicate that this protein has significant cryoprotective and heat-protective function via potential weak electrostatic binding to the target enzyme.

TH-3004_146106002

Table of Contents

TH-3004_146106002

Chapter 2: Experimental Techniques, Materials and Methods

TH-3004_146106002

Chapter 3: Generation of DHN1 tryptophan-cysteine mutants, protein expression, purification and characterization

Chapter 4: Probing the structural transitions in DHN1 protein by new tool:ProCharTS

TH-3004_146106002

TH-3004_146106002

Chapter 7: Summary and future perspectives

TH-3004_146106002

Chapter 1

Introduction and Review of Literature

TH-3004_146106002

1.1 Introduction:

The essential large molecules that are found in all living things, from the unicellular to the complex multicellular, can be sorted into four types− carbohydrates, proteins, lipids and nucleic acids. These macromolecules are essential for the survival and growth of living organisms. Among all the macromolecules, proteins are the most abundant (50% of the dry mass of cell of most cells) and diverse macromolecules present in the living organisms[1,2]. Proteins play an important role in sustaining cellular life through several essential processes like transport, cellular communication, storage, defense, structural support and speeding up the chemical reactions[3–5]. To understand the role of proteins in the cellular processes, their structural and functional insights are very important. To get better insights into proteins in cellular processes, investigation of their structure and function is important.

The techniques to study proteins have progressed significantly since 1958, when the three-dimensional structure of myoglobin was discovered using X-ray analysis[6]. At the earlier time, when a few crystal structures of protein were determined, it was found that some protein segments remained unresolved in electron density maps of X-ray crystallography and still participate in protein function[7–9]. Later in the 1990s, it was discovered that some of the segment that was not resolved in X-ray crystallography was unordered or disordered. Proteins with a partial disorder or fully disordered structure play a significant role in cellular functions have emerged as a new class of proteins recognized as Intrinsically Disordered Protein (IDP)/Intrinsically Disordered Region (IDRs).

1.2 Intrinsically Disordered Proteins (IDPs):

The structural and functional relationship has been associated entirely with the structure-function paradigm, where a unique amino acid sequence encodes a unique stable 3D protein structure for a specific biological function. This structure and function paradigm has dominated in biology for more than a century regarding a functional protein or a functional protein domain. In contrast to the structure-function paradigm, intrinsically flexible/disordered but biologically active proteins have gained the attention of the scientific community during the past two decades. These intrinsically flexible/disordered proteins are recognized by multiple terms such as floppy[10], rheomorphic[10], flexible[11],

TH-3004_146106002

natively denatured[12], natively unfolded[13,14], intrinsically unstructured[15,16], intrinsically disordered[17], malleable[18] and dancing protein[19], among several other terms. Amongst all the different names, over recent years, the term "intrinsically disordered protein" (IDP) is becoming more extensively used than other terms. Intrinsically Disordered Proteins (IDPs) gained attention due to their critical cellular role without any globular/3D structure[20–22]. They established a new regime based on the exhaustive analysis of available then literature data, the important hypothesis was formulated that naturally disordered proteins, instead of being just exceptions, represent a unique and very broad class of proteins[15,23–25].

IDPs are a class of proteins which exist as dynamic ensembles in which the atoms positions and backbone Ramachandran angle vary significantly over the time with no specific equilibrium values (unlike globular proteins), and these conformational ensembles involve non-cooperative conformational changes[26,27].

IDPs/IDRs remain unstructured in unbound form; however, in bound form with their binding partners like different membranes, metal ions, DNA, RNA, proteins, osmolytes and macromolecular crowding agents, they display the ability to gain structure. Many instances are reported even in the bound form; some portion of IDPs remains disordered[16,28–31]. In comparison with structures globular proteins and domains, IDPs/IDRs are different in their amino acid composition, hydrophobicity, net charge, sequence complexity, flexibility and degree of amino acid substitution over evolutionary time[17,32–34].

TH-3004_146106002

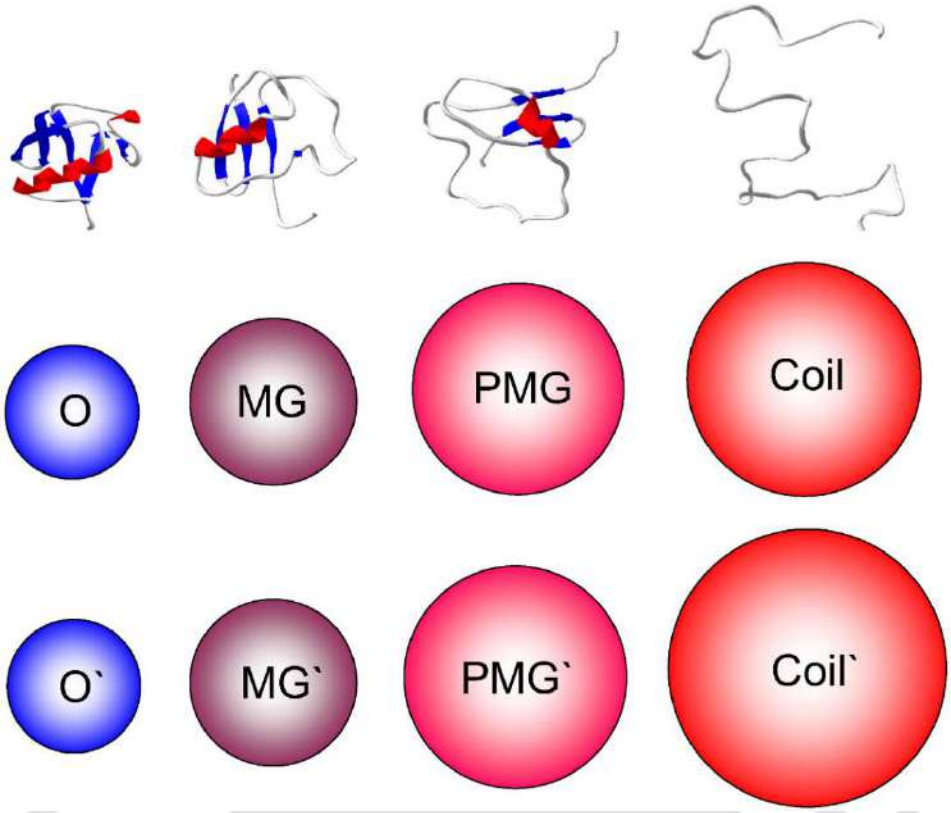


Figure 1.1 Illustrative example of Intrinsically Disordered Proteins. Collapsed (molten globule-like, MG) disorder; extended (pre-molten globule-like, PMG) disorder; and (coil-like, coil) disorder. Ordered globular protein of same length is shown for comparison. Figure represent model structure of a 100 residue-long polypeptide chain. **Middle line**: Relative hydrodynamic volume occupied by a 100 residue-long polypeptide chain in these four conformations. **Bottom line**: Relative hydrodynamic volumes occupied by a 500-residue polypeptide chain in these four conformations. Sphere in the middle and bottom line shows an increase in the hydrodynamic volume relative to the volume of the corresponding ordered protein. Adapted from: (Uversky and Dunker,2010b) [245]

This important class of protein increased curiosity in the scientific community, which may be attributed to several factors such as understanding the molecular mechanism by which they function, the transition from disorder to order, and their role in various significant processes[21,26]. IDPs are ubiquitous among all organisms, from single-cell organisms to complex organisms, plants and animals. Bioinformatics studies reveals that their prevalence increase with the increase in the complexity of the organisms[23]. An average of about 2 % or Archaeal, 4.2% of Eubacterial and 25% to 30% of Eukaryotic proteins are mostly

TH-3004_146106002

disordered[35,36]. Signaling proteins in Eukaryotes are 70% unstructured or have an intrinsically disorder region [37]. Despite being highly flexible and lacking stable secondary and tertiary structure IDPs are frequently involved in regulation, signaling and control pathways[38–40].

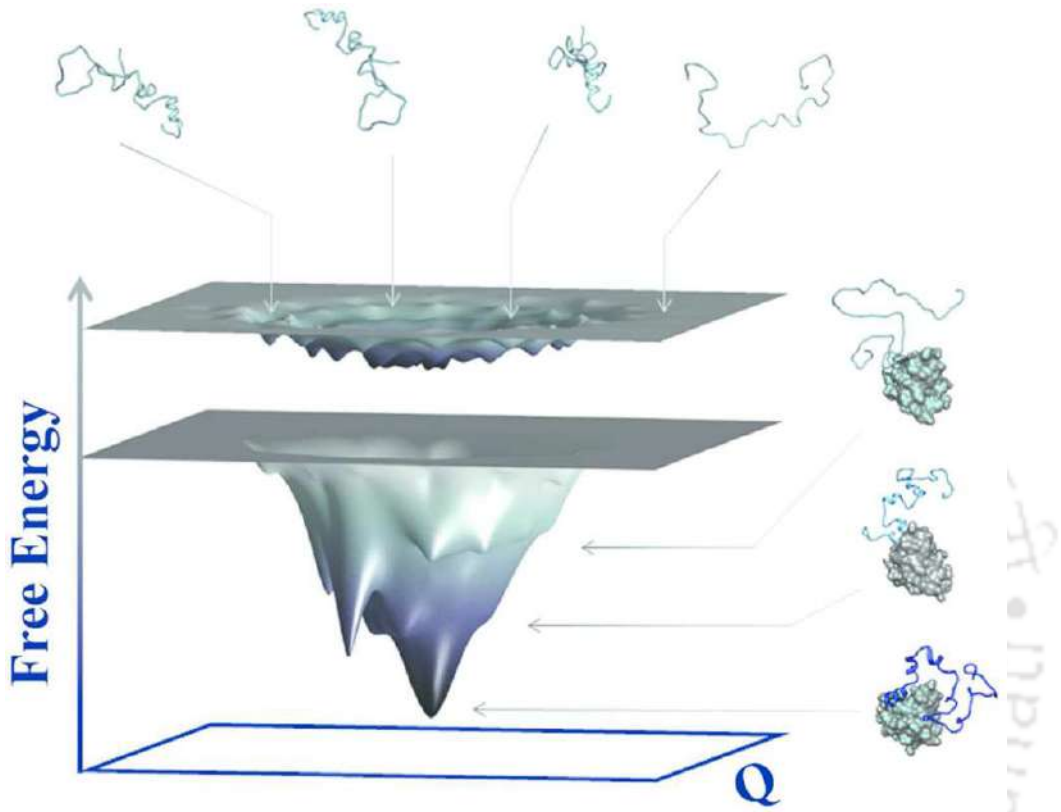


Figure 1.2 Free energy landscape hypothesis that can be generalized to describe IDPs. The free energy surface for an IDP in the unbound state can be characterized by a rugged but selectively flat landscape, whereas the IDP binding events funnel the free energy landscape, yielding well defining minimum. Adapted from: (Gibbs and Showalter) [246].

IDPs/IDRs can be also characterized by spatiotemporal heterogeneity, where parts of an IDP are disordered to varying degrees at one moment in time, but can change state at a future point in time [21]. Therefore, IDPs are not homogeneous, but represent a very complex mixture of a wide variety of partially foldable, potentially foldable, differently foldable, or completely unfoldable segments. **Figure 1.1** depicts the examples of intrinsically proteins and their hydrodynamic volume compared with ordered proteins. This behavior of an IDP is reflected in its free energy landscape with multiple local minima and without deep energy

TH-3004_146106002

minimum that can be found within ordered proteins. These simplified and shallow energy landscapes of IDPs make them extremely sensitive to numerous environmental changes that can further modify these complex landscapes. This spatiotemporal heterogeneity of energy landscapes describes the conformational plasticity of an IDPs/IDRs, where it interacts with various partners and subsequently fold in different ways. However, binding of IDPs to their partners is not always related with folding and other modes of binding includes highly dynamic or fuzzy complexes. **Figure 1.2** represents the free energy landscape of IDPs. Further, completely random coiled structures can only be achieved under harsh conditions where proteins have sufficient energies to overcome the ruggedness of the energy landscape. Under such conditions, the effective energy landscape is not as rugged. This might be the case at extreme high temperature, low pH, or high denaturant concentrations. Under mild denaturant concentrations or normal physiological conditions, rugged energy landscape imply uneven conformations distributions. Thus, disorder is not completely unstructured, flexible region. Although the molecules, globally or locally, might be unstable, for each of these there is still a preferred conformation. These conformations might have a low population time and, therefore, not be observed experimentally. Nevertheless, the preferred conformation, has a higher populations time than all alternate conformations.

1.2.1 Characteristics features of IDPs/IDRs:

1.2.1.1 Sequence signature of intrinsic disorder:

The propensity of intrinsic disorder is encoded in the uniqueness of amino acid sequences of IDPs. IDPs are significantly depleted in the bulky hydrophobic (Ile, Leu, Val and Met) and aromatic (Trp, Tyr and Phe) amino acids, which would be essential to form and stabilize the hydrophobic core of folded globular proteins, and also have low content of Cys and Asn residues. These in turn provide more solvent accessibility and less hydrophobic regions.

In contrast to folded globular proteins, IDPs are substantially enriched in polar and charged (Arg, Ser, Pro, Glu and Lys) with structure breaking (Gly and Pro) residues [25,32] [33]. These residues are called disorder-promoting amino acids. **Figure 1.3** shows the comparison of amino acid frequencies found in IDPs with the ordered proteins.

TH-3004_146106002

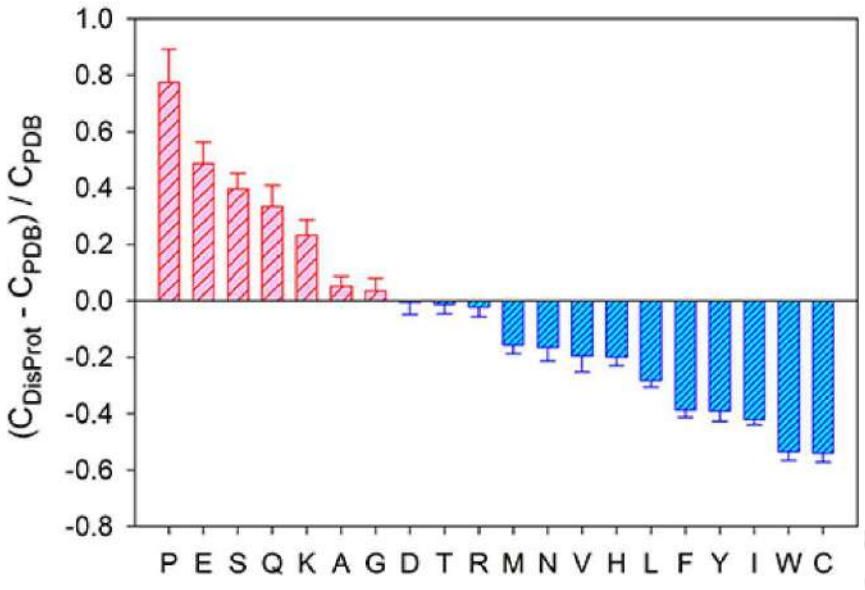


Figure 1.3 Amino acid determinants defining structural and functional differences between the ordered and intrinsically disordered proteins. Fractional differences in the amino acid composition between the typical IDPs from the DisProt database [247] and a set of completely ordered proteins [248] calculated for each amino acid residue. The fractional difference was evaluated as $(C_{DisProt}-C_{PDB}))/C_{PDB}$, where $C_{DisProt}$ is the content of a given amino acid in a DisProt database [247] and C_{PBD} is the corresponding content in the dataset of fully ordered proteins from PDB select 25 [248]. Adapted from: (Uversky,2019) [249].

1.2.1.2 Charge and Hydrophobicity:

One of the critical features of highly disordered proteins is their high net charges and low mean hydropathy. This is a prerequisite for the absence of compact structure in protein under physiological conditions[14,41]. Highly charged proteins can exhibit a prominent expansion at low ionic strength that correlated with their net charges. In aqueous solutions, numerous cases have been reported showing the formation of heterogeneous ensembles of collapsed structure from protein sequences rich in uncharged, polar amino acids and devoid of canonical hydrophobic residues[42–48]. **Figure 1.4** display the charge hydropathy plot for ordered versus disordered proteins.

TH-3004_146106002

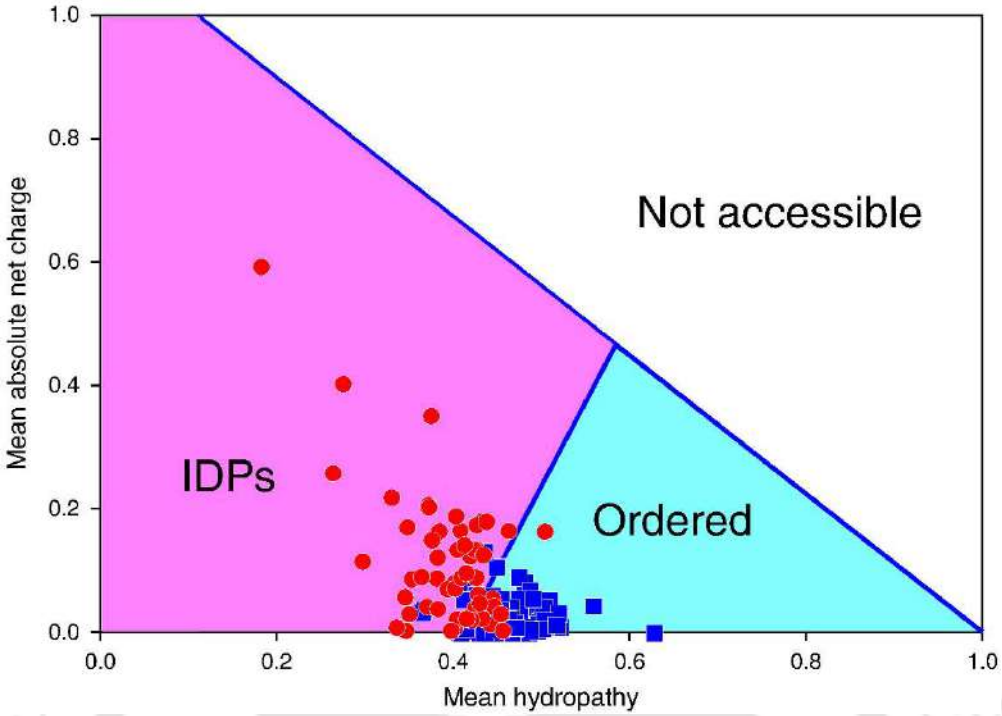


Figure 1.4 Charge-hydropathy plot for ordered proteins (blue squares) and IDPs (red circles), The area accessible to sequences encoding ordered and disordered proteins are shown in contrast colors, cyan and pink, respectively. These two areas are defined by two boundaries, the known boundary separating ordered proteins and extended IDPs and the boundary showing logical limits of the CH-space. Adapted from: (Uversky,2013) [71].

1.2.1.3 Binding promiscuity and plasticity:

One of the unique functional features of IDPs is their binding promiscuity and structural plasticity, the ability of one protein to bind to multiple partners and the ability to gain numerous binding-induced conformations [38]. Many IDPs can form highly stable complexes, or be involved in signaling interaction where they undergo constant "bound-unbound" transitions, thus acting as dynamics and sensitive "on-off" switches. The ability of these proteins to return to their highly dynamic and pliable conformation after the completion of a particular function, and their predisposition to gain different conformation depending on the peculiarities of their environment are unique properties of IDPs which allows them to exert different function in different cellular contexts according to a specific conformation state[26,49]. **Figure 1.5** illustrate how intrinsically disordered proteins are utilized by a hub of

TH-3004_146106002

proteins and the structural malleability of the intrinsically disordered region, which allows them to perform multiple functions.

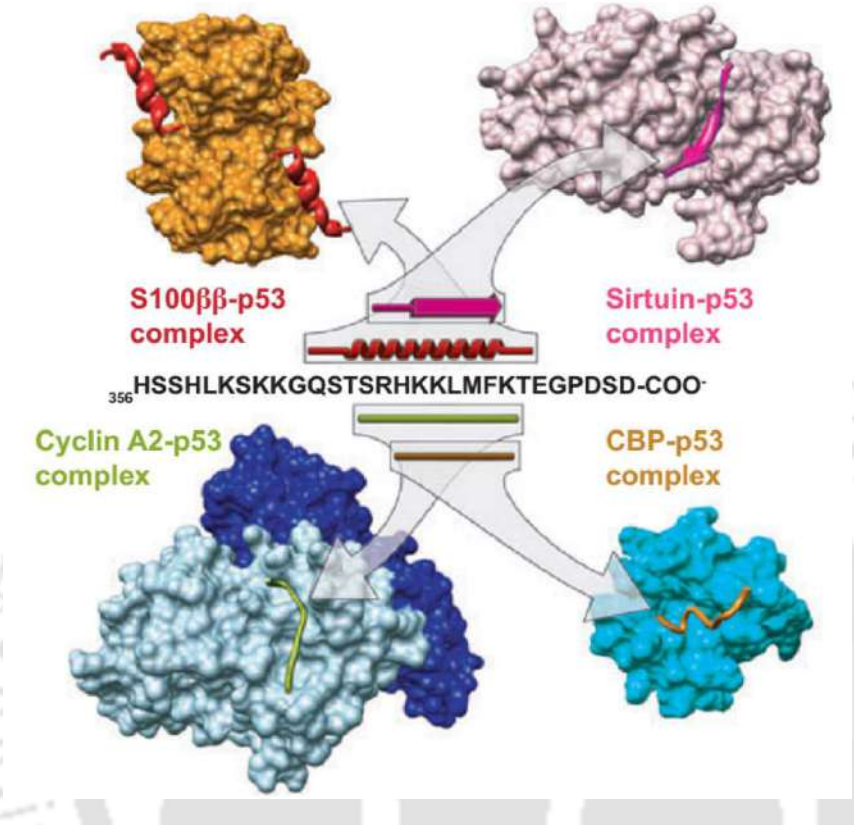


Figure 1.5 Sequence and structure comparison for the four overlapping complexes in the C-terminus of p53. The Primary, secondary and quaternary of p53 complexes. Adapted from: (Uversky,2009)[228].

1.2.1.4 Anomalous electrophoretic mobility:

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) is a well-established method generally utilized to evaluate the apparent molecular weight of proteins. IDPs have been known to show an anomalous behavior on SDS-PAGE. IDPs show larger apparent molecular weight compared to their actual molecular weight calculated from sequence or determined by mass spectroscopy. This anomalous behavior of IDPs is due to their unusually high content of charged amino acids and owing to less binding with SDS and causing slower migration on SDS-PAGE in contrast to the globular proteins[50–52]. It has been noticed that the molecular weight of IDPs determined by SDS-PAGE is generally

TH-3004_146106002

over-estimated by 1.2 -1.8 times higher than the actual one determined by mass spectrometry or theoretically from protein sequence[16].

1.2.1.5 Enhanced proteolytic sensitivity and heat resistance:

IDPs have been shown to have higher (5−7 fold) sensitivity to proteolytic cleavage compared to globular proteins[16,53,54]. This is due to the high flexibility and significantly less enzyme: substrate ratio (1:100−1:1000) required for rapid proteolysis of IDPs as compared to the globular proteins, which need the high (1:10−1:50) enzyme: substrate ratio[52,55]. Proteolytic recognition sequences in IDPs are more easily accessible and can adapt to the active site of protease due to their highly structure flexibility[56–58].

Numerous reports suggest that the IDPs are stable higher temperature i.e., stable to heat denaturation. This thermal resistance in IDPs is due to the low mean hydrophobicity and high net charges. This property is routinely used in the purification of IDPs[55,59,60].

1.2.2 Functional anthology IDPs/IDRs:

IDPs have a crucial role in several biological functions such as cell cycle control, transcriptional and translational regulation, stress tolerance, signaling, and disease[61–64]. Binding promiscuity and plasticity are the two key factors that confer functional advantages to IDPs. According to their mode of action, IDPs are classified into six general categories[16,65]. These categories are RNA and protein chaperone, display sites, effectors, assemblers, scavengers and entropic chains. **Figure 1.6** display the functional classification scheme of IDPs/IDRs. Among the six functional categories of IDPs, entropic chains function directly stem from their ensembles of structural states of similar conformational energies. In the other five classes, IDPs functions by molecular recognitions where they bind permanently (*effectors, assemblers, scavengers*) or transiently bind (*display sites and chaperone*) to the macromolecules or small ligands. Briefly, *Chaperones* function by evolving their disorder region to recognize, solubilize or loosen the structure of the misfolded ligand. A statistical analysis unveiled that the RNA chaperone has much higher disorder (40%) as compared to the protein chaperone (15%)[65]. *Display sites* are involved in the functions associated with post-translational modification by providing the flexible site

TH-3004_146106002

to the active site of modifying enzyme, which enables the transient but specific interaction[66,67].

IDPs that functions via permanent partner binding belongs either to *effectors, assemblers* or *scavenger*. *Effectors* function via binding and modifying (inhibition or activation) the partner enzyme. An example of a classical effector protein is $p21^{Cip1}$ and its homologue $p21^{Kip2}$ which has been shown to inhibit cyclin-dependent kinase (Cdks) and also to be able

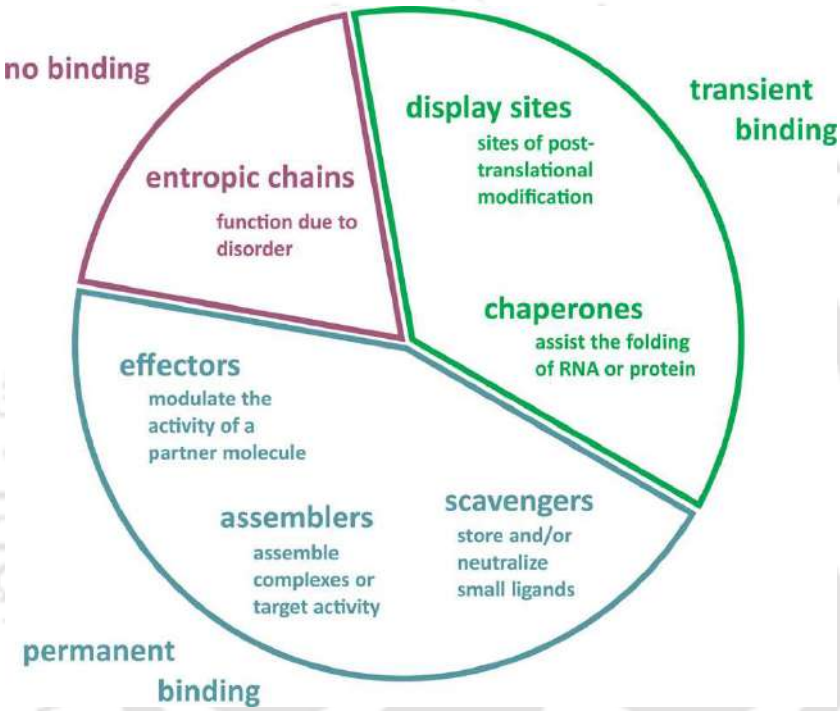


Figure 1.6 Functional Classification scheme of IDPs/IDRs. They bind either transiently as display post-translational modifications or as chaperone (green), or bind permanently as effector, assembler, or scavengers (dark-tone blue) Adapted from: (Van Der Lee et al., 2014)[250]

to assemble the cyclin-Cdk complex leading to Cdk activation[16,68]. Next, *Assemblers* are involved in the function to assemble multiple protein complexes like the ribosome, transcription preinitiation complex, cytoskeleton and chromatin or target the activity of the attached domain[69]. The last category, *Scavengers* functions by storing or neutralizing the small ligands via permanent binding to them. The classic example of a *scavenger* is casein(s), which prevents the calcium phosphate precipitation in the milk by capturing small seeds as they form[16].

TH-3004_146106002

1.2.3 Methodology used to characterize IDPs/IDRs:

As mentioned above IDPs/IDRs are either completely disordered or contain the long tail of an unstructured region; this makes them complex. It could be a mistake to analyse these proteins only by techniques developed for the characterization of the folded proteins[25,70–74]. IDPs/IDRs can be characterized by utilizing the following techniques mentioned below briefly.

X-ray crystallography: Even though X-ray crystallography is traditionally utilized for solving the three-dimensional structure of ordered proteins. It was noticed several times that some protein segment yields no discernible electron density, which may correspond to a disordered region. This missing electron density may arise from the flexible atom in IDPs, leading to non-coherent scattering[25,75]. Validation from other methods is required to confirm that the missing electron density is due to intrinsically ordered region or other technical issues[76].

NMR: Heteronuclear multidimensional nuclear magnetic resonance is the most versatile and powerful technique to investigate proteins disordered along their whole length. The high flexibility of the chain makes the chemical shift dispersion diminish and disorder easily recognizable. Additionally, recent advances permit complete assignment of resonance in the unfolded state and allow the sequence-specific identification of residual structure [77]. In addition, spin-relaxation and nuclear Overhauser effect measurement provides insights into internal molecular dynamics in the structured state. This technique is used to resolve the structure of numerous IDPs/IDRs[15,78–80].

Spectroscopy: The spectroscopic techniques most often used to characterize IDPs/IDRs are far-UV circular dichroism (CD) and fluorescence. The CD spectroscopy is used to determine secondary structure content in protein[54,8182]. The ellipticity spectrum of IDPs has a larger negative peak at around 200 nm and a value close to zero at 222 nm. A distinct CD spectrum from that of folded protein enables the identification of particularly or fully unstructured proteins. Further spectroscopic techniques, often combined with denaturing conditions, supplement these data. Fourier transform infrared spectroscopy also provides information on the secondary structure, whereas, near UV-CD and fluorescence

TH-3004_146106002

spectroscopy reports on the environments of aromatic amino acids (i.e., the lack of tightly packed hydrophobic core). Further, three-dimensional information across the IDPs can be determined using Förster Resonance Energy Transfer (FRET)[83,84]. Another important method that allows the characterization of secondary structure is Raman optical activity measurement[85]. This technique provides a snapshot of conformations and permits a distinction between dynamic and static disorder.

Hydrodynamics radius: Unfolded conformations of protein can also be detected and characterized by hydrodynamics techniques. Gel filtration (Size-exclusion chromatography), small-angle X-ray scattering[12], sedimentation analysis and dynamic light scattering provides the information on hydrodynamic parameter like Stokes radius and radius of gyration (R_G)[86,87]. In this state IDPs resembles the denatured state of globular proteins.

Some other independent techniques can add an extra dimension to these studies, Differential Scanning Calorimetry (DSC) unveils the absence of cooperative folding transitions characteristic of a well-defined tertiary fold. The extreme proteolytic sensitivity of IDPs results from proteases cleavage substrates at site that are sterically accessible and flexible enough to make productive contact with the enzyme. Furthermore, SDS-PAGE can be used to preliminary identification of IDPs/IDRs, where IDPs/IDRs shows anomalous behavior and show apparent M_w 1.2-1.8 times higher than the real one calculated from the sequence data or measured by mass spectrometry.

1.2.4 Intrinsically disordered proteins in plants:

Plants are exposed to several adverse environmental conditions including both biotic and abiotic stresses. Water deficit condition are inflicted upon drought, cold and salinity. They have developed numerous adaptations in response to such adverse conditions. In such adaptation to stresses IDPs play an important role in plants. The plethora of reports indicates that intrinsic disorder is critical for plant protein function in various responsive, regulatory and signaling processes[38,88–90]. The protein disorder in plants follows the same trends reported for other species. **Figure 1.7** displays the distribution of intrinsically disordered proteins in plants. As an example, a genome-wide analysis of protein disorder in

TH-3004_146106002

Arabidopsis thaliana showed that the biological processes more enriched in disordered proteins were related to cell cycle, signaling, DNA metabolism, RNA splicing etc.[91].

The dehydrins are the most studied examples of disordered plant proteins[52,92–94]. This large and diverse family of proteins, involved in response to the drought and other environmental stresses, is almost entirely disordered. This family includes protein chaperones such as ERD10 and ERD 14[52,95] as well as proteins involved in the binding of metal ions, protection of membrane, and global protection of cell during the highly compact dry state characteristic of plant seeds[92].

Another plant-specific family of proteins heavily relies on disorder for functioning in the GRAS family[92]. These proteins play an essential role in plant development and are involved in signal transduction cascades; they act as integrators to hormone response. **Figure 1.8** illustrates the role of intrinsic disorder in five plant protein family**.**

Recent studies in plants show different classes of protein that are upregulated in various abiotic stressful conditions. One such protein family is Late Embryogenesis Abundant proteins (LEA), LEA proteins are well-known to be upregulated in abiotic stress conditions such as water desiccation, low temperature, and high salinity[32].

TH-3004_146106002

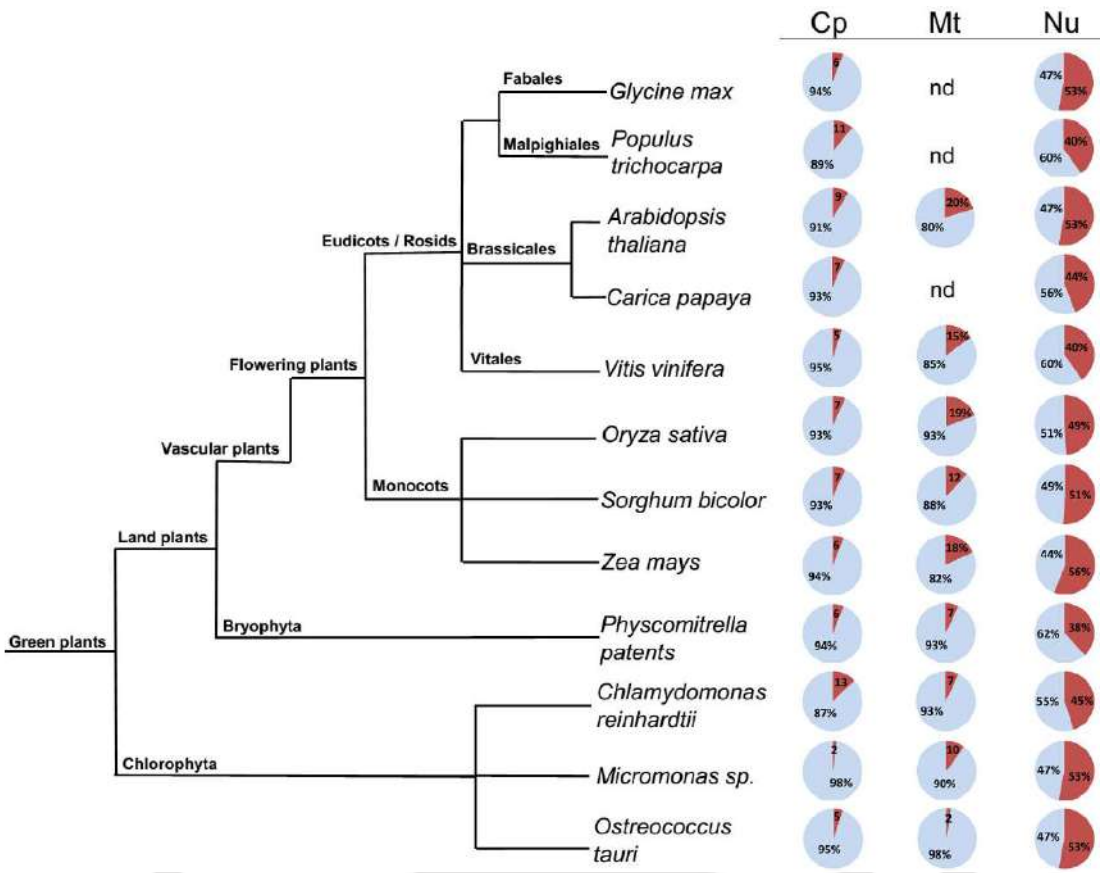


Figure 1.7 Distribution of predicted discovered segment with L ≥ 30 in plants. Disordered in chloroplast (Cp), mitochondria (Mt), and nuclear (Nu) proteomes are shown. Percentage of intrinsically disordered proteins are in red and percentage of non-disordered proteins are in light blue. Adapted from: (Yruela and Contreras-Moreira, 2012)[251].

TH-3004_146106002

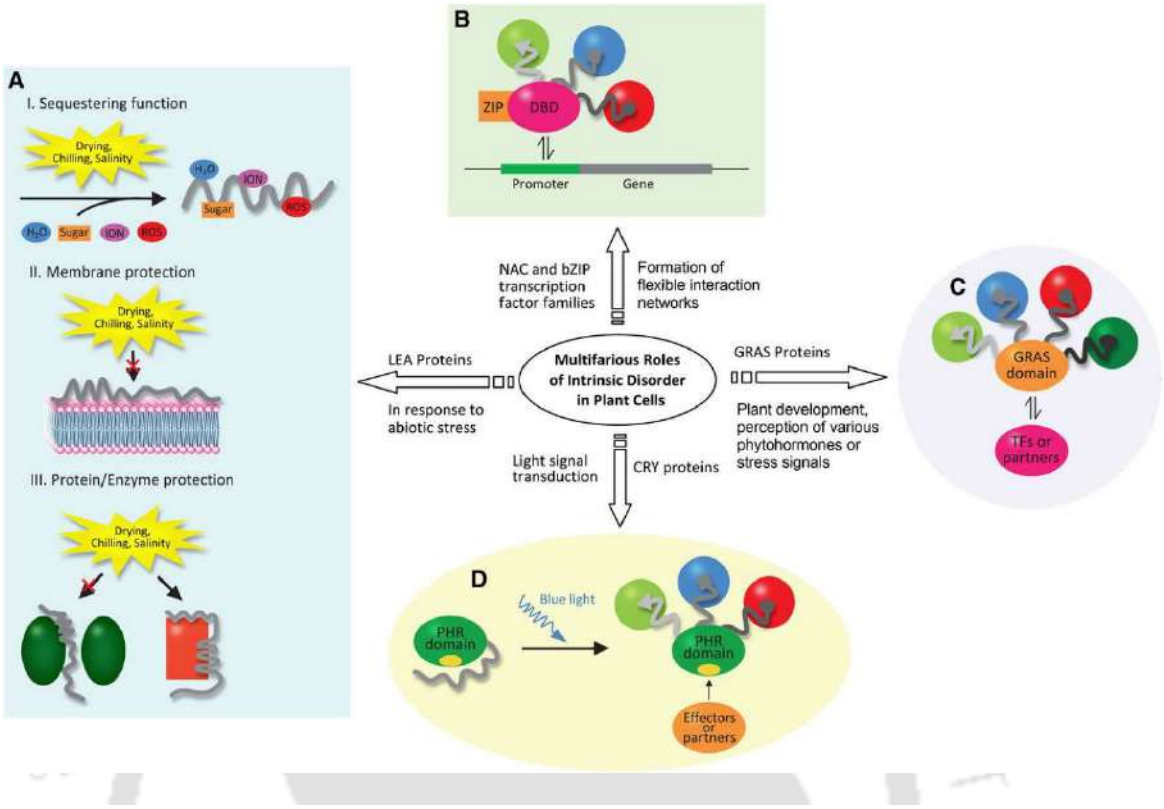


Figure 1.8 Multifarious Roles of Intrinsic Disorder in five plant protein Families. (**A**) Intrinsically disordered LEA proteins (gray coil) protect plant cells against abiotic stress through sequestering functions. (**B**) The disordered TRD of a TF (gray coil) helps from flexible transcriptional regulatory networks by interacting with various partners (spheres in different colors) under different conditions. (**C**) Disordered N-domain of GRAS proteins control plant development by interacting with various partners (spheres in different colors) for perception of phytohormones and environmental signals, which coordinate interaction between the GRAS domain and TFs or protein partners. (**D**) Disordered C-terminal domain of CRYs act as light-dependent switches that transduce the signal to a specific protein-protein interaction between the C-terminal domain of a CRY and partner (Spheres in different colors) or between the PHR of a CRY and an effector/partner, which initiates the photomorphogenic program. The yellow oval in the PHR represents an active site masked by a disordered C-terminal domain when in darkness. Adapted from**:** (Sun et al.,2013) [92]**.**

TH-3004_146106002

1.2.5 Dehydrin:

Dehydrins (dehydration proteins) are a member of the group 2 LEA protein family[97–99]. Dehydrins are found in different intracellular locations in the nucleus and cytoplasm. Other locations include mitochondria, chloroplast and plasma membrane. This diverse localization of dehydrins is reflected in its diverse functions. These include preventing electrolyte leakage across the cold stressed membrane, preventing lipid peroxidation, anti-aggregation, and cryoprotection[100–102].

By definition, a dehydrin contains at least one copy of the K-segment[103]. K-segment is 15 amino acid, Lys-rich motif (EKKGIMDKIKEKLPG) that can be found in 1-11 copies[104]. The S-segment, which contains 5-7 Ser residues, is not necessarily present in all dehydrins. Another, Y-segment (V/T)D(E/Q)YGNP is present in one or two copies. The Φ-segments are less conserved motifs and interspersed between all conserved segments[105]. Dehydrins are classified into five subclasses based on the conserved segments, i.e., SK_n, K_n, Y_nSK_n, K_nS Y_nK_n[106]. Recently, another conserved segment was identified and named as F-segment[107] based on the pair of hydrophobic F residue in the core of the sequence.

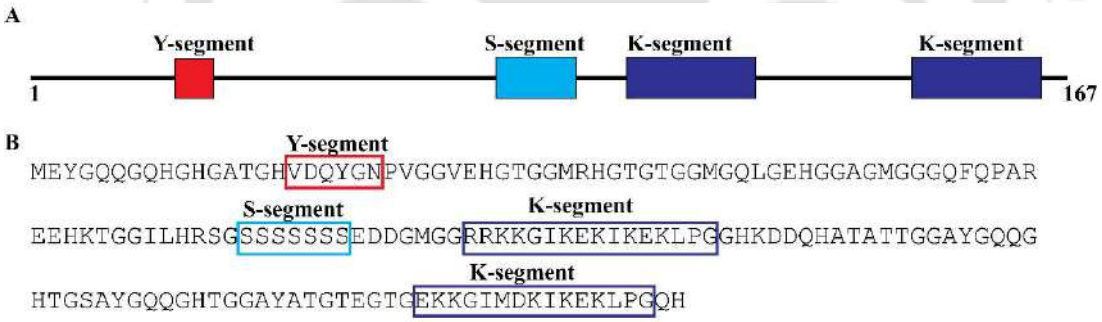


Figure 1.9 Schematic representation of DHN1 protein of *Zea mays*. (**A**) Schematic representation of DHN1 protein contains conserved Y, S and K segments in red, cyan, blue color, respectively. (**B**) The amino acid sequence showing the conserved segments red, cyan, blue color in boxes.

Figure 1.9 shows the schematic illustration of Y, S and K segments in DHN1 protein of *Zea mays*. The K-segment can form an amphiphilic α-helix. This segment in dehydrins may help to stabilize the protein and the cellular membrane under the environmental stress

TH-3004_146106002

conditions. The *Citus unshiu* K3S type dehydrin CUCOR19 was found to form an α-helix in the presence of SDS (Sodium dodecyl sulfate) [101]. Similarly, the K-segment in maize dehydrin is essential to bind to the anionic phospholipids vesicles[108]. The K-segment of wheat dehydrin DHN-5 is required for the protection of B-glucosidase and lactate dehydrogenase (LDH) activities *in vitro*[109]. The cold-induced dehydrin LTI30 can bind to the cellular membrane via its K-segment, which locally folded into α-helix on the cellular membrane surface. The Lti30's K-segment electrostatically trap the negatively charged lipid head[110].

The Y-segment of dehydrins has sequence similarity to the nucleotide-binding site of plant and bacterial chaperones. However, there is no experimental evidence suggesting that the Y-segment bind nucleotides[103,111]. The S-segments is suggested to promote dehydrin interaction with the specific signal peptide followed by their translocation into the nucleus[103,112,113].

The role of conserved segments in dehydrins are is not yet fully understood. Dehydrins which contain more than one K-segment tends to be expressed during cold-stress, whereas the Y-segment dehydrin tends to be correlated with desiccation and salt stress[101]. To better understand the physiological role of dehydrins, several *in vitro* assays have been performed. One of the extensively performed assays is the enzyme cryoprotection, where the commonly used enzyme is Lactate dehydrogenase[94]. Dehydrin also gains structure in the presence of SDS micelles and liposomes [114]. The mechanism by which the dehydrin protects the membrane and crucial enzymes from desiccation and cold stresses needs further investigation.

Dehydrin can protect nucleic acids in plant cells during seed maturation and stress responses by interacting electrostatically with vesicles of both zwitterionic and negatively charged phospholipids[94]. Many studies have reported that dehydrins may play a protective function on enzyme or phospholipid as molecular chaperone or molecular shield[111,115]. Dehydrin ERD10 and ERD14 have chaperone activities with rather wide substrate specificity and interact with phospholipid vesicles through electrostatic force under cold stress, while dehydrin *PpDHNA* and *PpDHNB* from *Physcomitrella patens* provide protection as molecular shield to lactate dehydrogenase (LDH) under osmotic and freezing conditions[116].

TH-3004_146106002

KS-type dehydrin (AtHIRD11) can recover the LDH activity denatured by Cu^{2+} in *Arabidopsis* [117].

Hughes and Graether[111] proposed that the dehydrin would stay preferentially localized near the enzyme rather than bind with LDH. Furthermore, there was an extremely weak association between the K-segment of K_2 and LDH, which was mediated by long range-electrostatic forces. Kovacs et al[52] argued that the strict distinction between classical chaperone action and molecular shield activity might not be tenable. Although the protective function of dehydrins to plants exposed to abiotic stress conditions in vivo or in *vitro* have been reported, the protective mechanism is not completely understood.

To study IDPs/IDRs spectroscopically, the intrinsic chromophores of proteins like Tryptophan, Tyrosine and peptide bond were exploited. Brief discussion of the conventional and non-conventional chromophores with focus on recent developments under novel approaches to investigate proteins spectroscopically are mentioned in next section.

1.3 Chromophores in Proteins:

Molecules undergo electronic transitions such that their electrons are promoted from lower to higher molecular orbitals when excited by higher energy electromagnetic radiation in the range of UV (200−400 nm) and visible (400−800 nm) spectrum. The phenomenon of UV light absorption by proteins has been utilized to study the structural characteristics of macromolecules from the early days of molecular biology[118]. The strong absorption of water below 170 nm restricts the absorption studies of macromolecules below 170 nm. There are three classes of chromophores in proteins that are predominantly utilized for absorption studies.

1.3.1 Peptide Bond:

In proteins, the most frequent electronic transitions include the peptide bond which has characteristic absorbance in the far-UV region. Predominantly, the π- electron of peptide bonds are delocalized over the N, C and O atoms and non-bonding n-orbital electron near the O atom. These electronic transitions show two distinct peaks: one strong peak at 190 nm with ε= 7000 $M^{-1}cm^{-1}$ (arise due to π-π* transitions) whereas another peak of weaker

TH-3004_146106002

intensity at about 210-200 nm with ε= 100 $M^{-1}cm^{-1}$ (arise due to n-π* transitions)[119,120.] Additionally, the third peak around 175 nm involving an n-σ* transitions can be observed. The secondary structural changes in the protein can cause alteration in the peptide bond absorption. Both poly-L-glutamic acid and poly-L-lysine shows changes in absorption intensities with conformational changes[121].

1.3.2 Aromatic amino acids:

In apoprotein, other than peptide bond which shows the absorbance below 220 nm, the chromophores which has absorption properties are the side chain of aromatic amino acids namely tryptophan (Trp), tyrosine (Tyr) and phenylalanine (Phe). These amino acids display significant absorption at wavelength longer than 230 nm (in near UV-region) and are most frequently utilized as a gold standard for quantification and estimation of proteins. Among the three aromatic amino acids, Trp shows the strongest absorption (ε= 5600 $M^{-1}cm^{-1}$) arising from the indole group present in the side chain[122]. It principally comprises two major peaks; the first around 220 nm (ε= 36000 $M^{-1}cm^{-1}$) and the second around 280 nm (ε= 5600 $M^{-1}cm^{-1}$)[123]. Next, tyrosine is another aromatic amino acid with significant absorption properties at 275 nm (ε= 1400 $M^{-1}cm^{-1}$) and at 222 nm (ε= 9000 $M^{-1}cm^{-1}$)[124]. The optical properties of tyrosine residue in proteins is highly pH sensitive and most among the aromatic amino acids[125]. Phe displays the two absorption bands: one at around 257 nm (ε= 200 $M^{-1}cm^{-1}$) and another at around 205 nm (ε= 9600 $M^{-1}cm^{-1}$). Change in pH does not influence Phe's absorption[124]. Additionally, the imidazole group in the side chain of Histidine, Cysteine and methionine also shows the appreciable absorbance in lower wavelength region between 185-200 nm and 230-240 nm range, respectively[126]. However, the disulphides (cystine) show the absorption at longer wavelengths, between 250-270 nm (ε= 300 $M^{-1}cm^{-1}$) as compared to cysteine[127].

1.3.3 Charge transfer Complexes:

Aside from the chromophores discussed above, another category of chromophores relies upon Charge Transfer (CT) of electron from donor to acceptor. Although such charge transfer complexes are relevant in various biomolecules like DNA, lipids, vitamins, porphyrins, saccharides and even in microorganism[128], here we highlight only prominent

TH-3004_146106002

cases of Charge Transfer investigated in protein system. Numerous processes like photosynthesis in microorganisms and plants[129], and electron transport chain in respiratory pathway relies on the formation of charge transfer complexes. The two species of microorganism, *Rhodobacter sphaeroides* and *Rhodopseudomonas viridis* (purple bacteria), were commonly studied where the CT transitions occur between two bacterial chlorophylls. The most common investigated examples of biomolecules are the enzymes and metalloproteins. Typically, metal-ligand complexes in metalloproteins shows the absorption through charge transfer transitions[130]. The prosthetic group (e.g., heme, flavin, carotenoid) form complexes with proteins to perform important functions. Such metal-protein complex (for e.g., Azurin, Xanthine oxidase) along with many important coenzymes of proteins such as FAD, NADH and NAD^+ display strong absorption in UV-Vis region. FAD absorbs at 450 nm (ε= 11300 $M^{-1}cm^{-1}$)[131], NADH at 340 nm (ε= 6220 $M^{-1}cm^{-1}$) and NAD^+ at 259 nm (ε= 16900 $M^{-1}cm^{-1}$)[132].

Further, charge transfer (CT) mechanism has also been investigated in peptides containing aromatic amino acids, Trp and Tyr [133], where the conversion of indolyl into phenoxy radical is mediated via electron transfer either through single step super-exchanges or by multistep hopping mechanism. Moreover, long range electron transfer by the amide bond through the α-helical peptides as intermediate relay junctions has also been reported [134].

1.3.4 Unconventional Chromophores:

Most of the chromophores mentioned above that are either present in apoprotein like peptides and aromatic amino acids or in metal complexes. In recent time numerous reports have been published that reveal the unusual UV-Visible absorbance spectra beyond 315 nm and sometimes extend till 800 nm, in aqueous solution of charged amnio acids S/R-[135], L-Lysine-HCl[136], poly-L-lysine-HCl[137]. Similar absorbance has been found in various crystals of amino acids[138], and other solutes[139], but the chromophore behind such unusual absorption was not known. Furthermore, the luminescence from protein as well as the from various biopolymer, nucleic acid, and carbohydrate has been a recent topic of interest. Multiple reports of intrinsic fluorescence and deep-blue luminescence arising from oligomeric proteins[140], amyloid aggregates[141], crystals[142]. Moreover, the dendrimers (PAMAM) and

TH-3004_146106002

various small molecular structures, natural and synthetic polymers, that are devoid of any known conventional fluorophores, also displayed the intrinsic blue luminescence.

1.4 Protein Charge Transfer Spectra (ProCharTS):

1.4.1 ProCharTS Absorption:

In 2001 our research group reported that aqueous solution of L-Lysine.HCl[136] at high concentration displays absorption (270-350 nm) and fluorescence emission, but the origin behind this unusual spectra remained mysterious. Other research groups also observed and reported similar observations in amino acids[138,139]. This type of unusual absorbance in the near UV-Visible region also seen in the small molecules such as HCl and NaOH[143]. To unveil the reason behind the unusual electronic absorption of L-Lysine in amino acids and proteins, a synthetic protein, α_3C was investigated. This protein is rich in charged amino acids (Lys and Glu) and devoid of aromatic amino acid residues. The investigation displayed absorption beyond 300 nm [144], as shown in **Figure 1.10 B,** similar to the previously reported proteins rich in charged content[137], **Figure 1.10A**. Both computational (Time-dependent density functional theory (TDDFT) calculations on MD simulation snapshots) and experimental investigation elucidated that various Lys-Glu, Lys-Lys, and

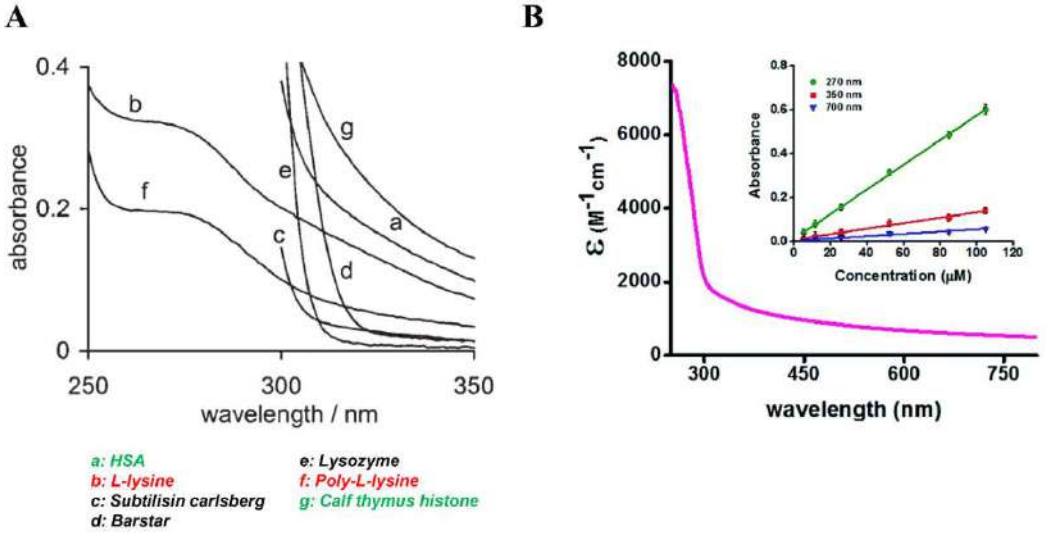


Figure 1.10 ProCharTS absorption in (**A**) charged rich proteins and (**B**) α_3C protein. Adapted from**: (**Homchaudhuri and Swaminathan, 2004; Prasad et al., 2017)[137,144].

TH-3004_146106002

Glu-Glu contact in the protein to be the reason for such unusual UV-Vis absorption spectra in the range of 250-800 nm [144]. This phenomenon is named as **Protein Charge Transfer Spectra (ProCharTS)**.

ProCharTS absorption arises due to photoinduced charge transfer from: polypeptide backbone to NH_3^+ in the lysine; COO^- in glutamate to polypeptide backbone; and COO^- in glutamate to NH_3^+ in lysine. The ProCharTS absorption intensities are dependent on 3D spatial proximity between charges in Lys-Lys, Glu-Glu and Lys-Glu side chain head groups across the proteins as shown in **Figure 1.11**. Further, a computational study shows the similar phenomenon in naturally charged amino acids (Lys, Glu, Arg, Asp and His) and phosphorylated amino acids (Tyr, Thr and Ser) [145].

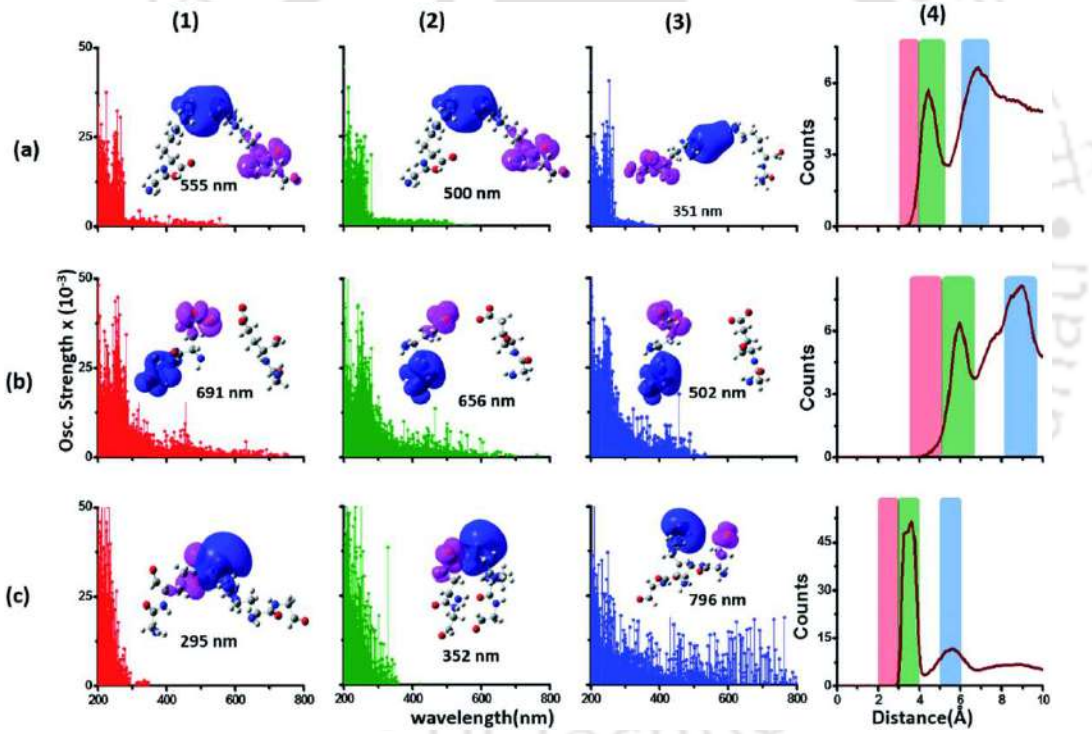


Figure 1.11 Simulated absorption spectra (Oscillator strength vs wavelength) for dimers of **a**) Lys−Lys, **b**) Glu−Glu, and **c**) Lys−Glu. Subfigures **1**,**2**,**3** represents the Charge Transfer at different distances; Subfigure **4** represents their respective Radial Distributions. Adapted from: (Prasad et al., 2017)[144].

TH-3004_146106002

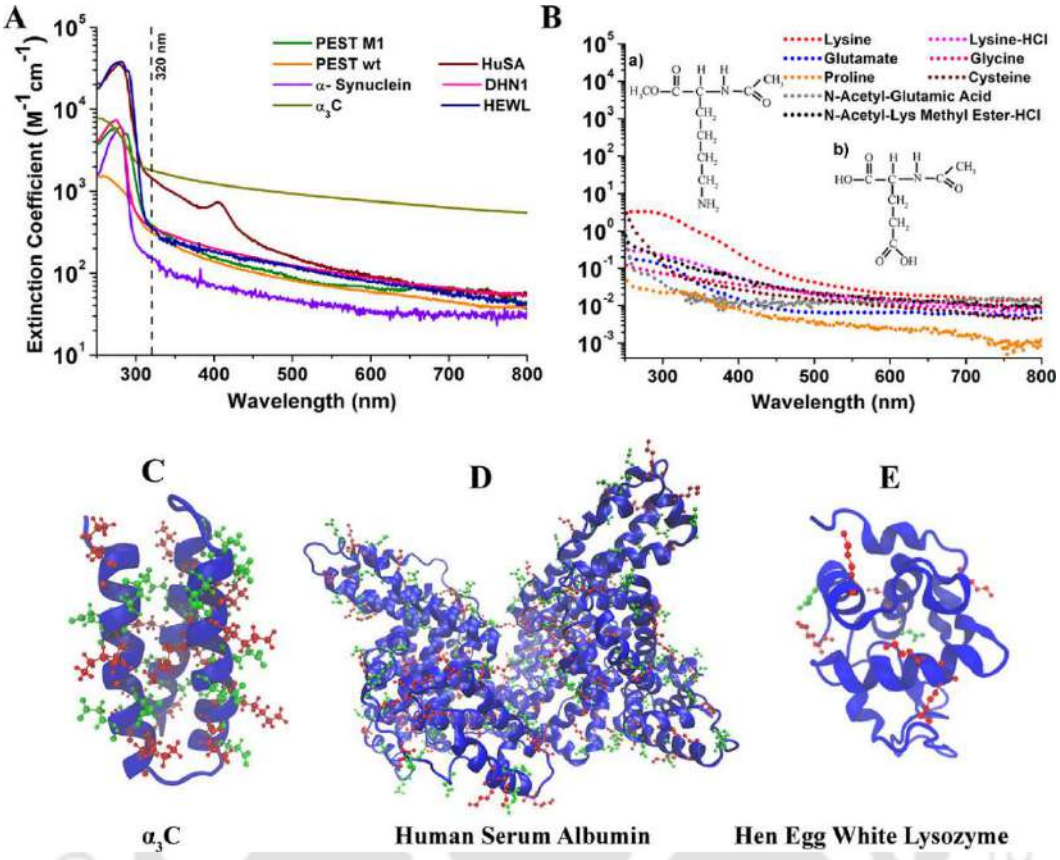


Figure 1.12 ProCharTS absorption spectra and extinction coefficient of (A) charged proteins and **(B)**charged amino acids. Subfigures in **C**, **D** and **E** highlights the charged content Lys (red) and Glu (Green) in α_3C., Human Serum Albumin (HAS), Hen egg White Lysozyme (HEWL), respectively. Adapted from: (Kumar et al., 2020) [146].

1.4.2 ProCharTS luminescence:

Following absorption, the photochemical features of intrinsic novel luminescence in proteins were investigated by several groups. Numerous reports on the intrinsic deep-blue blue luminescence in proteins, amyloid aggregates, and even crystals of amino acids and proteins hint at different hypothesis. Same as intrinsic absorbance, the origin of unusual luminescence is quite unclear. An investigation made on different monomeric proteins enriched in charged amino acids and devoid of any chromophores displayed the luminescence characteristics[146], as shown in **Figure 1.12**. The monomeric nature of protein was confirmed by the linear increase in luminescence with increase in protein concentration

TH-3004_146106002

and any possibility of oligomeric or aggregated state which contributes to luminescence was ruled out, in contrast to earlier reports that suggested only aggregation or oligomerization gives such blue luminescence[147,148]. Additionally, the concentrated aqueous HCl and NaOH solutions have been reported to show such intrinsic luminescence where no aggregated moieties were present[143].

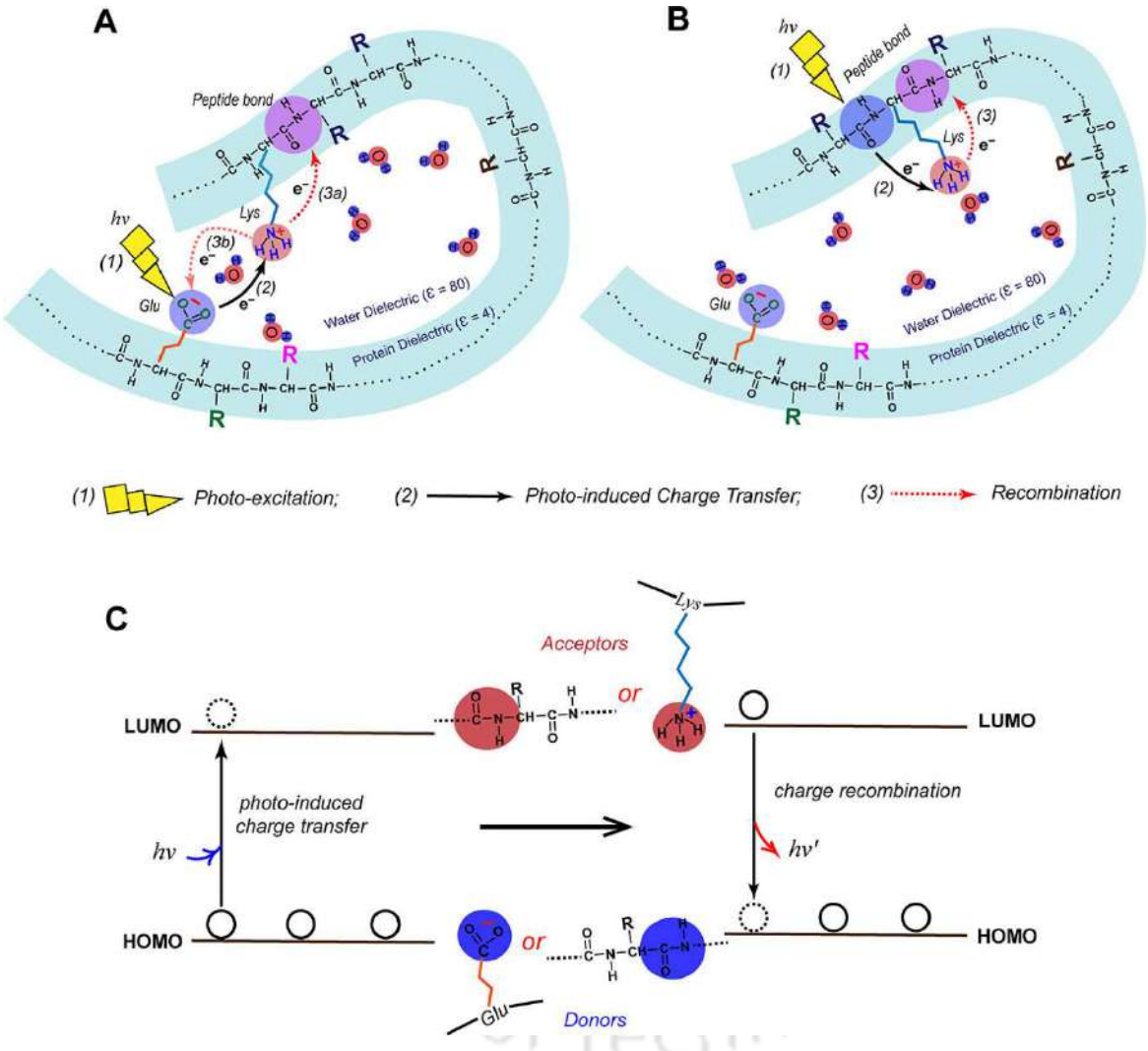


Figure 1.13 Schematic illustration showing the pathway for photoinduced charge transfer from (**A**) Glu to Lys and (**B**) peptide backbone to Lys, followed by charge recombination. Panel (**C**) shows the schematic of photoinduced charge transfer and charge recombination luminescence. Adapted from: (Kumar et al.,2020) [146]**.**

Like the ProCharTS absorbance, ProCharTS luminescence was proposed to be the origin of such novel luminescence from monomeric proteins and amino acids[146]. The concept of

TH-3004_146106002

charge transfer transitions between various charged species present in proteins is therefore raised to be the reason for blue luminescence. This novel luminescence is suggested to arise from charged recombination phenomenon between excited state electrons and the holes in the proteins. Charge recombination is proposed to be mediated by a) direct electronic recombination to the hole or b) delayed recombination with electrons hopping on the protein backbone and then recombination either with the same or a new hole, as shown in **Figure 1.13.**

ProCharTS in monomeric proteins is generally characterized by low quantum yield and multiple luminescence lifetimes. The origin of multiple luminescence lifetimes where charge recombination is mediated via numerous pathways also supports the charge recombination hypothesis. ProCharTS luminescence was found to be highly correlated with the product of the molar extinction coefficient and luminescence quantum yield[146]. Thus, it was found that ProCharTS luminescence too is highly dependent on the close spatial 3D proximity of the head groups of charged amino acids and is more pronounced in charge-rich proteins compared to the proteins with low or moderate charge content.

Since the dehydrin protein is enriched in charged amino acid residues and devoid of bulky and hydrophobic amino acids, it would be interesting to investigate this protein using the newly discovered tool: Protein Charge Transfer Absorption Spectra (ProCharTS). This method relies on the charged amino acid content of proteins.

In the present thesis, I have attempted to investigate the structural transitions of 167 amino acid long DHN1 (YSK_2 type) protein from *Zea mays* by a newly discovered ProCharTS and other biophysical techniques. Based on the literature we design the following objective for the thesis work:

TH-3004_146106002

1.5 Objectives for the thesis work:

1. Generation of DHN1 tryptophan−cysteine mutants, protein expression, purification and characterization.
2. Unravelling the structural transitions in DHN1 by the new tool: ProCharTS absorption.
3. Luminescence characteristics of ProCharTS in DHN1.
4. Investigation of structural dynamics and functional characterization of DHN1 and its mutants.

The DNA encoding for the DHN1 protein (of *Zea mays*) was cloned into the pET19b expression vector between *BamHI* and *NcoI* restriction site. The cloning was carried out by Suzanne G. Jepson and Timothy J. Close [149] and the clone was kindly gifted to our lab.

In objective 1, the substitution mutations at chosen site in DHN1 were done using the site-directed mutagenesis (SDM) to incorporate the single tryptophan and single tryptophan−cysteine pair. The purpose behind the incorporation of tryptophan and tryptophan−cysteine pair in DHN1 protein was to investigate the folding properties of DHN1 protein in the presence of binding partners (here, we used the SDS, a negatively charged molecule that mimics the lipid bilayer) and the role of N-terminal in DHN1 folding. The DHN1 and its mutant DHN1 CW1 and DHN1 W3 were successfully expressed and purified using boiling lysis and affinity chromatography methods. Subsequently, the sequence-based characterization of DHN1 and its mutants DHN1 CW1 and DHN1 W3 was employed by amino acid composition analysis, disorder prediction plots, charge hydropathy plots and hydrophobic cluster analysis. These all tools served as to predict the intrinsic disorder propensity in DHN1 and its mutants. Further, spectroscopic techniques like absorption, fluorescence and circular dichroism were used to characterize the DHN1 and its mutants.

As previously discussed, IDPs have abundant content of charged amino acids in their sequence and are devoid of aromatic amino acids like tryptophan/tyrosine. Aromatic amino acid serves as the intrinsic spectral probe to structure analysis using various techniques like

TH-3004_146106002

absorption and fluorescence spectroscopy. Thus, their absence in IDPs limits its structural investigation to a few techniques like circular dichroism and NMR. Moreover, to investigate the DHN1 protein structural transitions and folding properties, we have to incorporate the tryptophan, and tryptophan−cysteine pair in DHN1 using SDM. Further, to determine the 3D distance across the protein using FRET measurement, an additional external probe (acceptor) is required. This imposes severe restrictions on monitoring the structural transitions IDPs in a label-free approach using UV-Visible spectroscopy. Thus, there is a demanding need of alternate intrinsic chromophores to study the structure and dynamics of IDPs. Here, in objective 2, we employed a newly discovered tool: ProCharTS absorption to probe the structural transitions induced by various factors such as change in temperature, pH, and salt. This method is quite simple and inexpensive to study the conformational changes in IDPs as compared to other conventional methods.

In next objective 3, we examined the weak intrinsic ProCharTS luminescence in DHN1 protein. The characterization of emissive properties of DHN1 ProCharTS at spectral range of 320-410 nm was done by using the steady-state fluorescence measurement and time-resolved fluorescence spectroscopy. Further, the influence of weak intrinsic ProCharTS luminescence on the tryptophan lifetime in DHN1 mutant; DHN1 CW1 was investigated.

In final objective 4, we explored the structural transitions and functional role of DHN1 and its mutants DHN1 CW1 and DHN1 W3. The structural dynamics of DHN1 protein in the presence of TFE and heat induced transitions were studied. Further, we looked at the DHN1 protein folding in the presence of SDS and determined the 3D distances between the incorporated tryptophan and cysteine using FRET measurement. Cryoprotection and heat-protection activity measurement of DHN1 and its mutants DHN1 CW1 and DHN1 W3 was carried out by employing the LDH activity assay.

TH-3004_146106002

Chapter 2

Experimental Techniques, Materials and Methods

TH-3004_146106002

2.2 Experimental Techniques:

The principles of some important biophysical techniques employed in this study are explained here in brief.

2.1.1 Site-directed mutagenesis:

In vitro site-directed mutagenesis (SDM) has been established as an important tool to investigate protein dynamics and the structure-function relationship. There are several approaches for the SDM using the Polymerase Chain Reaction (PCR)[150], but these methods generally utilize single-stranded DNA (ssDNA) as the template[151,152] This can be quite labour intensive or technically complicated. One of the broadly used methods is QuikChange[152]. Here, we used the QuikChange II site Directed Mutagenesis kit (Agilent Technologies), which allows site-specific mutation in virtually any double-stranded plasmid. The conventional process required DNA subcloning, ssDNA rescue and a specialized vector. Moreover, the unique restriction site, multiple transformations or in vitro methylation treatment steps make the process complex. The use of the QuikChange II site Directed Mutagenesis kit simplifies the overall process. This includes the quick three-step procedure for generating mutants with greater than 80% efficiency in a single reaction. The simple protocol utilizes either miniprep plasmid DNA or cesium-chloride-purified DNA. The basic procedure uses a supercoiled double-stranded DNA (dsDNA) vector with an insert of interest and two synthetic oligonucleotide primers 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′-Gm6ATC-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]. As a result, only mutation-containing synthesized DNA

TH-3004_146106002

remains in its intact from[151]. Further, nicked vector DNA containing the desired mutation was transformed into XL1-Blue super-competent cells or DH5α competent cells. The incorporation of the desired mutation at chosen site in the plasmid vector of interest can be confirmed by isolation of the plasmid and subsequent DNA sequencing.

2.1.2 Spectroscopy:

Spectroscopy is defined as the study of the interaction of light with matter[153]. Light is electromagnetic radiation (ER) of oscillating electric and magnetic fields in phases perpendicular to the propagation direction and with respect to each other, as shown in **Figure 2.1**. Electromagnetic radiation can disturb the distribution of charges and spins within a molecule upon exposure to light. This perturbation could lead to changes in the electrical and magnetic properties of the molecules[154]. These changes in molecules can be detected from the properties of the radiation that emerges from the sample. The simplest property that one can measure is the fraction of incident light absorbed or dissipated by the sample.

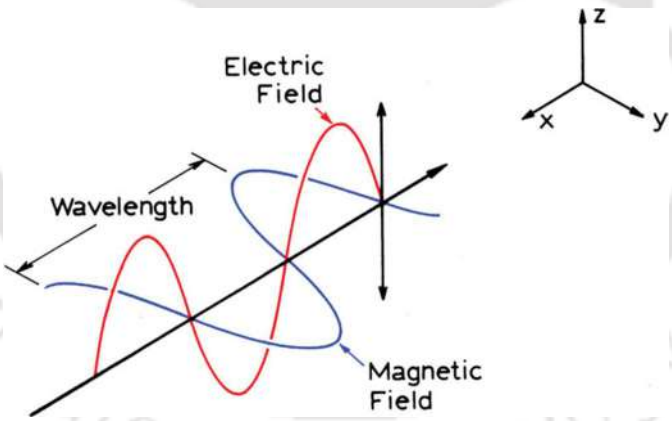


Figure 2.1 Illustration of vertically polarized light. Adapted from Principle of Fluorescence Spectroscopy[156] by J. R. Lakowicz, 3rd edition, 2006.

Techniques like optical absorption spectroscopy depend exclusively on this measurement. Aside from this, the radiation emitted by the sample at another wavelength different from the excitation wavelength can also be detected, upon which the techniques like fluorescence, phosphorescence and Raman scattering are based. Additionally, complex techniques like fluorescence polarization and circular dichroism (CD) are based on the type/order and degree of polarization of the emitted radiation from the sample[155]

TH-3004_146106002

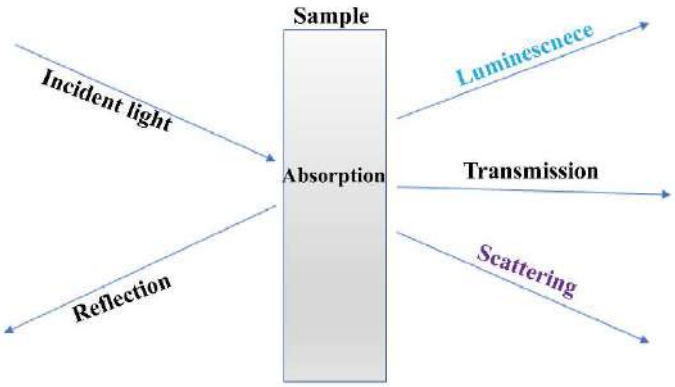


Figure 2.2 Illustration of light matter interaction.

Table 2.1. Biologically useful spectroscopic regions Adapted from: Cantor and Schimmel[155] 1980.

Typical Wavelength (cm)	Approx. Energy (kcal mole^{-1})	Spectroscopic Region	Techniques and Application
1×10^{-11}	3×10^{8}	γ-Rays	Mössbauer
1×10^{-8}	3×10^{5}	X-rays	X-ray diffraction, scattering
1×10^{-5}	3×10^{2}	Vacuum UV	Electronic spectra
3×10^{-5}	1×10^{2}	Near UV	Electronic spectra
6×10^{-5}	5×10^{3}	Visible	Electronic spectra
1×10^{-3}	3×10^{0}	IR	Vibrational spectra
1×10^{-2}	3×10^{-1}	Far IR	Vibrational spectra
1×10^{-1}	3×10^{-2}	Microwave	Rotational spectra
1×10^{0}	3×10^{-3}	Microwave	Electronic paramagnetic resonance
1×10^{1}	3×10^{-4}	Radio Frequency	Nuclear magnetic Resonance

There is a very broad range of wavelengths utilized in the spectroscopy of biological molecules. The list of the biologically useful spectral regions and associated techniques is enlisted in **Table 2.1**. In the current work, we mainly used absorption, fluorescence and circular dichroism spectroscopy.

TH-3004_146106002

2.1.2.1 Absorption:

Absorption spectroscopy in biology mainly deals with the absorption among the biomolecules in the visible or ultraviolet region. The absorption process of a photon (smallest possible unit of light) occurs very fast, in the time scale of 10^{-15} s [156]. Through the absorption of a photon, the electronic structure of the molecules is usually shifted from an electronic ground state into an energetically higher electronic excited state[157]. In the excited state, the three-dimensional electronic probability densities are different from the ground and all other state probability densities. Only photons having energies corresponding to the energy difference between these electronic states can be

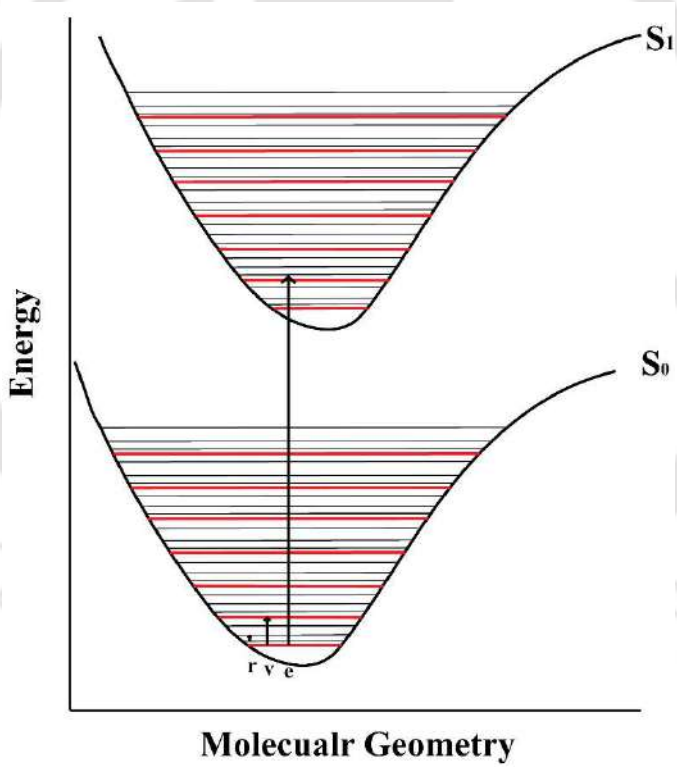


Figure 2.3 Different energy levels and transition between S_0 and S_1. Transitions r, v and e corresponds to rotational, vibrational and electronic levels, respectively. Adapted from Cantor and Schimmel[155] 1980.

absorbed by the molecules. This is the basis of colour of a compound: only light of a particular wavelength is absorbed by the compound. The light of other wavelengths is either transmitted or scattered. The Photon energy corresponding to the absorbed wavelength equals the necessary excitation energy for the particular transitions.

TH-3004_146106002

2.1.2.1.1 Basics of UV-Visible spectroscopy:

When electromagnetic radiation impinges on the molecule, there is an induction of dipole within the system. This induction of dipole can be expressed as

$$\mu_{ind} = \tilde{\alpha}.E \qquad 2.1$$

Where, μ_{ind} represent the induced dipole moment, $\tilde{\alpha}$ denotes the polarizability of the molecules, and *E* represents the electric field of the light. In the above equation, both *E* and μ_{ind} are variable that fluctuates with time.

Absorption of light leads to the change in the energy state of molecules. If a molecule is initially existing in state A will change to state B upon interaction with light. The probability of transitions from state A to state B is described by the transition dipole moment (μ_{ba}). μ_{ba} can be calculated by integration $\int \psi_b^* \tilde{\mu} \psi_a dt$ and can also be expressed as $\langle \psi_b | \tilde{\mu} | \psi_a \rangle$. Here the wavefunction of two state A and B is represented as ψ_a and ψ_b, respectively. The wavefunction in the presence of light can be represented as

$$\psi(t) = C_a(t)\psi_a e^{-iE_a t/h} + C_b(t)\psi_b e^{-iE_b t/h} \qquad 2.2$$

Here, the probabilities of the system, whether it will exist in state A and B relate to coefficients C_a *and* C_b. The energies of states A and B are represented by E_a *and* E_b, respectively.

Hamiltonian in the presence of light can be written as

$$\hat{H}' = \hat{H} + V(t) \qquad 2.3$$

Here *V(t)* is the effect of light on the system. This interaction energy between light and a molecule can be given by

$$\mathrm{V(t)} = \tilde{\mu}.E_0 e^{i\omega t} \qquad 2.4$$

The probability of the system existing in state B at time *t* can be written as

TH-3004_146106002

$$P_b = |C_b(t)|^2 \quad 2.5$$

The absorption intensity can be computed by the rate at which energy is taken up by the molecules from the incident light beam. The transition rate dP_b/dt can be written as

$$dP_b/dt = B_{ab}I(\nu) \quad 2.6$$

where, B_{ab} denotes the transition rate per unit energy density of the radiation and $I(\nu)$ represent the energy density incident on the molecules at frequency ν. B_{ab} can be expressed as

$$B_{ab} = (2/3)(\pi/h^2)|\langle\psi_b|\tilde{\mu}|\psi_a\rangle|^2 \quad 2.7$$

The molar extinction coefficient is one of the significant factors in light absorption. When light with intensity I_0, impinges on the sample molecules, in a layer perpendicular to the direction of light propagation sufficiently thin *(dl),* the fraction of incident light absorbed by sample molecules can be written as

$$-dI/I = C\acute{\varepsilon}dl \quad 2.8$$

Where, *I* represent the intensity of transmitted light, *l* denotes the path length in cm travelled by the light in solution, *C* is the molar concentration of the absorbing molecules, and $\acute{\varepsilon}$ signify a proportionality constant called the molar extinction coefficient ($M^{-1}cm^{-1}$). Upon integration of equation 2.8 for the entire sample, we find,

$$\ln(I/I) = C\varepsilon(\lambda)l \quad 2.9$$

Upon conversion of equation 2.9 to log base 10, we have the common form of the Beer-lambert Law:

$$A(\lambda) \equiv \log(I_0/I) = C\varepsilon(\lambda)l \quad 2.10$$

Where, $\varepsilon = \varepsilon'/2.303$, and A is called the absorbance or the optical density.

The molar extinction coefficient is correlated to the Einstein coefficient for stimulated absorption as

$$B_{ab} = (1{,}000c/N_0h)\int(\varepsilon'/\nu)d\nu \quad 2.11$$

TH-3004_146106002

Here, N_0 denotes the Avogadro's number, and c is the velocity of light with frequency ν

By using equation 2.11 and equation 2.7, and converting ε' to ε, we obtain the subsequent relation

$$D_{ab} \equiv |\langle\psi_b|\tilde{\mu}|\psi_a\rangle|^2 = 9.180\, x\, 10^{-3} \int (\varepsilon/\nu) d\nu \ \ (debye)^2 \qquad 2.12$$

Where, D_{ab} is called the dipole strength. D_{ab} can be determined by integrating the area under an absorption band (band of interest should be resolved sufficiently from nearby bands). After determining the D_{ab}, one can measure the length of the transition dipole.

One more significant measure is the oscillator strength, f_{ab}, which compares the intensity of absorption to that expected from a three-dimensional harmonic oscillator. This can be expressed as

$$f_{ab} = (8\pi^2 mc/3h\nu) D_{ab} = 4.315 \times 10^{-9} \int \varepsilon(\nu) d\nu \,(\text{dimensionless}) \qquad 2.13$$

Where, m represents the mass of the electron. f_{ab} can have the value in the range of 0.1 to 1, for a strongly allowed transition.

2.1.2.2 Fluorescence Spectroscopy:

Fluorescence spectroscopy has been used as one of the primary research tools in biochemistry and biophysics. The fluorescence technique is vastly used in clinical chemistry, DNA sequencing, and genetic analysis by fluorescence in situ hybridization (FISH) and many other areas. The primary reason behind its vast uses is its sensitivity and ease of doing.

Fluorescence is a type of luminescence. Luminescence is defined as the spontaneous emission of radiation from an electronically excited species or from a vibrationally excited species which are not in thermal equilibrium with its environment[158].

TH-3004_146106002

Subsequent to light absorption which takes places in femtosecond (10^{-15} s) time-scale, many processes usually occur. This comprises the excitation of fluorophore to some higher vibrational level of either S_1 or S_2. With some exceptions, molecules in condensed phases quickly relax to the lowest vibrational level of S_1. This rapid relaxation called the internal conversion, occurs in picoseconds (10^{-12} s) or less, which is generally completed prior to emission. In this process, the loss in energy is caused by vibrational relaxation or collision with solvent or other molecules. This process of internal conversion is temperature dependent, and an increase in temperature can cause a decrease in fluorescence by dissipating the lost energy into solvent as heat. The internal conversion process points out that all the excited-state electrons are relaxed to the lowest vibrational level of S_1 before the beginning of any fluorescence emission. Emission of photons accompanying the S_1 to S_0 relaxation is called fluorescence (10^{-9} s).

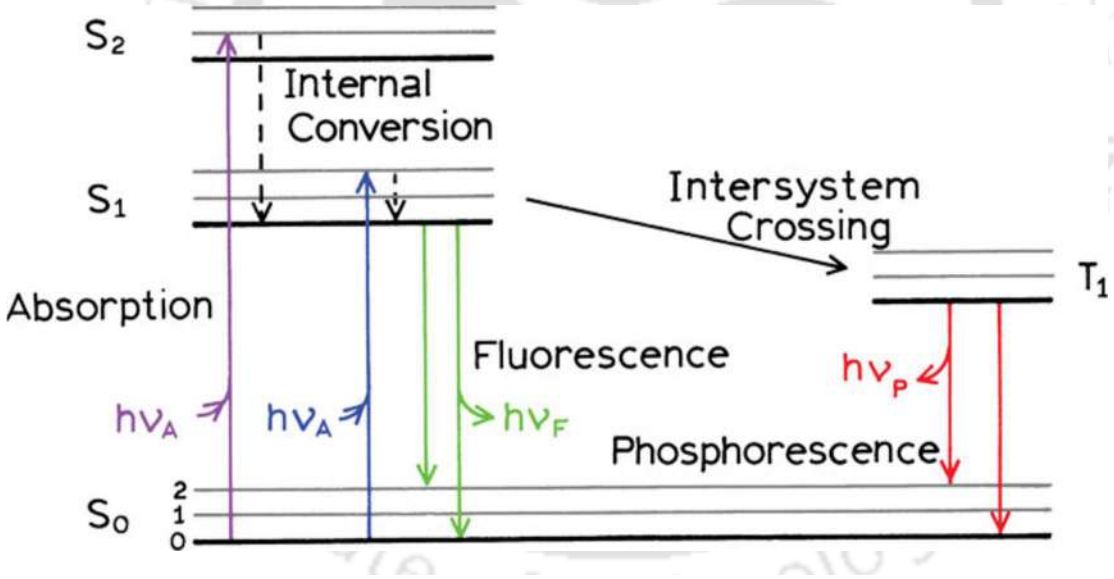


Figure 2.4 Simplified form of Jablonski Diagram. Adapted from Principle of Fluorescence Spectroscopy[156] by J. R. Lakowicz, 3rd edition, 2006.

Molecules in the excited state S_1 can also undergo a spin conversion to the first triplet state, T_1, which is termed as the intersystem crossing. The subsequent emission from the T_1 state is called as the phosphorescence (~10^{-3} s) and is shifted towards the longer wavelengths. The rate constant of for triplet state emission is several orders of magnitude smaller compared to the fluorescence; this is because of the spin-forbidden transition from T_1 to the

TH-3004_146106002

singlet ground state. The presence of heavy atoms like iodide and bromine favours the intersystem crossing by spin orbit coupling.

Aside from the internal conversion and intersystem crossing, the rate of fluorescence emission also may be affected by quenching. Any process that causes a decrease in the fluorescence intensity of a fluorophore is known as fluorescence quenching. This can occur by multiple mechanisms, either by depopulating the excited state fluorophores or by forming a non-fluorescent complex with the ground state fluorophore.

Different molecules like oxygen, amines, halogens, and electro-deficient molecules like acrylamide can acts as collisional quenchers, and the process is diffusion limited. The quenching process can be denoted as

$$S_b + Q \overset{k_q}{\rightarrow} S_a + Q + heat \qquad 2.14$$

Where, S_a is the ground state energy level and S_b is the excited state energy level of a molecule, Q is the quencher, and k_q represent the biomolecular quenching constant.

All these non-radiative processes (internal conversion and intersystem crossing), along with quenching, decrease the rate of fluorescence. The observed fluorescence lifetime (τ) can be defined as

$$\tau = \frac{1}{(k_r + k_{isc} + k_{ic} + k_q[Q])} \qquad 2.15$$

Where, k_r, k_{isc}, k_{ic} and k_q represents the rate of decay of fluorescence, intersystem crossing, internal conversion and fluorescence quenching, respectively. [Q] denotes the molar concentration of quencher. The above equation can be simplified as

$$\tau = \frac{1}{(k_r + k_{nr})} \qquad 2.16$$

k_r *and* k_{nr} signifies the rate constant for radiative and all non-radiative decay, respectively. Fluorescence lifetime in the absence of any non-radiative process is called intrinsic or natural lifetime (τ_n), represented by

TH-3004_146106002

$$\tau_n = \frac{1}{k_r} \qquad 2.17$$

The fluorescence efficiency or the quantum yield (ϕ) is defined as

$$\phi = \frac{k_r}{k_r + k_{isc} + k_{ic} + k_q[Q]} = \frac{k_r}{k_r + k_{nr}} \qquad 2.18$$

The intrinsic lifetime (τ_n) can be also be calculated from the measured lifetime *(τ)* using the following equation

$$\tau_n = \frac{\tau}{\phi} \; or \; \phi = \frac{\tau}{\tau_n} \qquad 2.19$$

The above equation illustrates the relationship between the quantum yield and lifetime. This shows the decrease in the quantum yield reflected as the decrease in fluorescence lifetime.

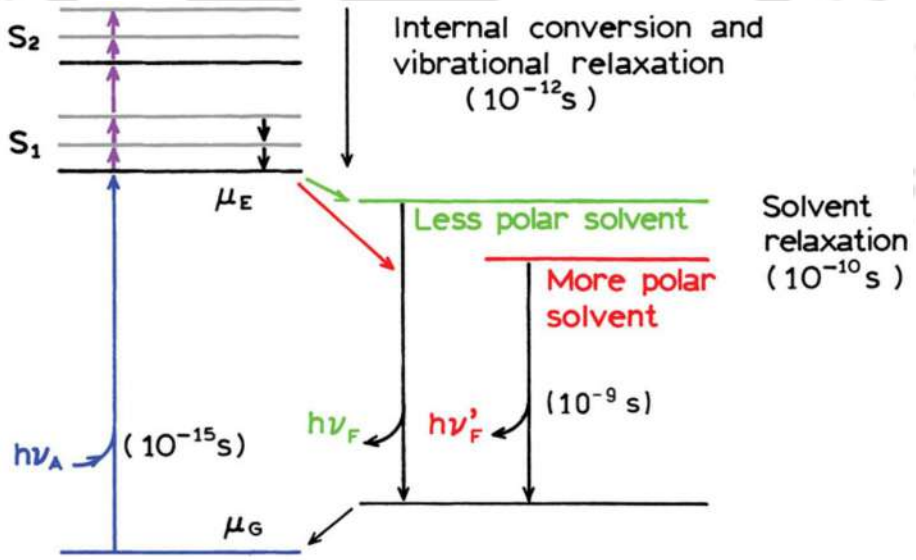


Figure 2.5 Jablonski diagram showing the effect of solvent relaxation on the energy of excited state. Adapted from Principle of Fluorescence Spectroscopy[156] by J. R. Lakowicz, 3rd edition, 2006.

In addition to the above-mentioned modes of non-radiative decay, polar solvent molecules like water can also affect the fluorescence emission. If the solute and solvent molecules possess a dipole moment, the solute-solvent interaction results not only from these

TH-3004_146106002

permanent dipole moments, but also from the polarizability of the molecules[158]. The relatively longer timescales of fluorescence emission avail ample time for solvent molecules to interact and reorient around the excited state of a fluorophore. Mostly, the fluorophores have a larger dipole moment in the excited state (μ_E) compared to their dipole moment in the ground state (μ_G). Subsequent to excitation of the fluorophores, the solvent dipoles can reorient or relax around the excited-state dipole of the fluorophore, (μ_E). This reorientation or relaxation of solvent dipoles around the excited-state dipole of fluorophores, lowers the energy of the excited state, as illustrated in **Figure 2.5**. This process is called solvent relaxation, and this occurs in a time scale of ~1-10 ps.

As shown in **Figure 2.5**, the lowering of the energy of the excited state causes the emission at longer wavelengths (or at lower energies). Such a shift in emission bands can be induced by a change in the solvent nature or composition. Thus, solvent relaxation results in the observed Stokes shift in the fluorescence spectra of any fluorophore.

The Stokes shift is the difference between the absorption maxima and the emission maxima of the fluorophore, and can be denoted as

$$Stokes\ shift, \Delta\bar{\nu} = \bar{\nu}_A - \bar{\nu}_F\ (cm^{-1}) \qquad 2.20$$

Here, $\bar{\nu}_A\ and\ \bar{\nu}_F$ are the wavenumber (cm^{-1}) of absorption and emission, respectively. Upon the increase in solvent polarity, this becomes larger and can result in a significant Stokes shift. The polar fluorophores are more sensitive to solvent polarity. The change in energy between the excited state and ground state because of solvent relaxation is defined by Lippert-Mataga equation as,

$$\bar{\nu}_A - \bar{\nu}_F = \frac{2}{hc}\left(\frac{\varepsilon - 1}{2\varepsilon + 1} - \frac{n^2 - 1}{2n^2 + 1}\right)\frac{(\mu_E - \mu_G)^2}{\alpha^3} + constant \qquad 2.21$$

Where, h represents the Planck's constant, c denotes the speed of light and α represent the radius of the cavity in which the fluorophore resides. $\bar{\nu}_A\ and\ \bar{\nu}_F$ are the wavenumber (cm^{-1}) of absorption and emission, respectively. While n is the refractive index and ϵ is the dielectric constant of the medium.

TH-3004_146106002

It should be noted that in Lippert-Mataga equation, the term $(\varepsilon - 1/2\varepsilon + 1)$ accounts for the spectral shift due to the reorientation of the solvent dipoles and the redistribution of the electron in the solvent molecules. Another term $(n^2 - 1)/(2n^2 + 1)$ accounts only for the redistribution of electrons. The difference between these two terms accounts for the spectral shift due to the reorientation of solvent molecules, and hence termed as the orientation polarizability (Δf), i.e.

$$\Delta f = \left(\frac{\varepsilon - 1}{2\varepsilon + 1} - \frac{n^2 - 1}{2n^2 + 1}\right) \qquad 2.22$$

However, the constant term in the equation 2.21 accounts for the minimal losses at vibration levels of excited state that is dissipated as the heat in the solvent. This suggests that even in most non-polar solvents, the absorption and emission maxima cannot overlap.

Fluorescence signals from the excited sample can be measured by two modes, either by steady state fluorescence or time-resolved fluorescence techniques.

2.1.2.3 Steady-state fluorescence:

In steady-state fluorescence, the sample is continuously illuminated with a continuous beam of light at a specific wavelength, and subsequently the fluorescence emitted by the fluorophores is recorded as a function of a wavelength. In this method, there is an equilibrium between the population of ground state molecules with the excited state molecules. This method can be used to calculated the quantum yield of fluorophores and measurement of different fluorescence properties like steady state anisotropy, Förster resonance energy transfer (FRET), quenching and many more. This kind of measurement provides time-averaged parameters of fluorophores in excited state. The variation in fluorescence intensity and emission maxima can yield valuable information about the solvent exposure and degree of polarity around the fluorophore. All this information provided by the steady-state fluorescence can be utilized to study several biological phenomena.

TH-3004_146106002

2.1.2.4 Steady-State Fluorescence Anisotropy:

Fluorescence anisotropy is the most sensitive way to detect the size and shape of molecules in solution. This reveals information about the size, shape and rigidity of various molecular environments. Anisotropy is used to explore protein-protein interaction, protein folding, membrane fluidity and protein aggregation. In fluorescence anisotropy, the sample is excited with the vertically polarized light (along Z axis) 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 weighted over the intensity decay of the fluorophore and divulges the information about the overall rotational motion of molecules.

$$r_{ss} = \frac{\int_0^\infty r(t)I(t)dt}{\int_0^\infty I(t)\,dt} \qquad 2.23$$

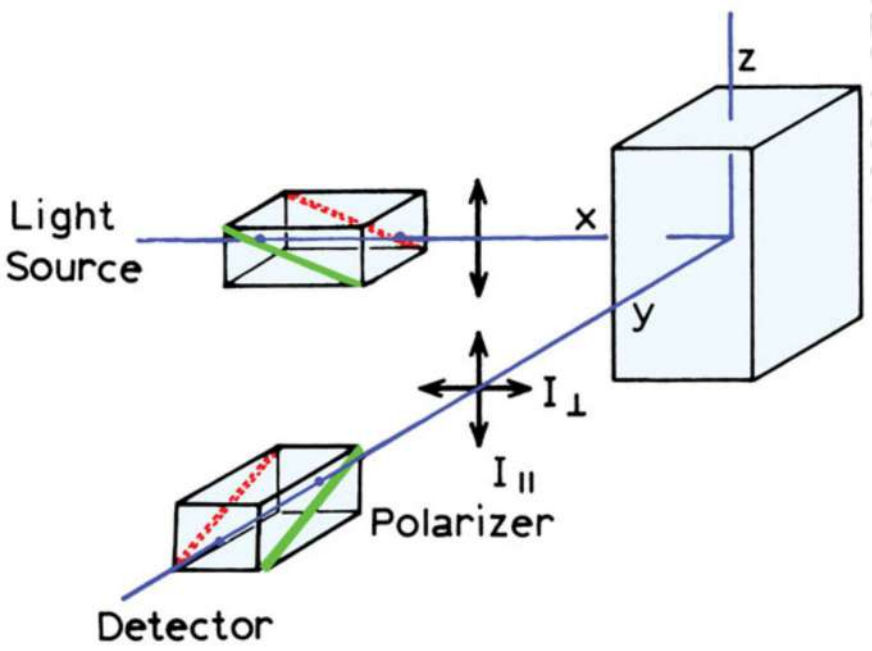


Figure 2.6 Illustration of schematic of Fluorescence anisotropy measurement. Adapted from Principle of Fluorescence Spectroscopy[156] by J. R. Lakowicz, 3rd edition, 2006.

TH-3004_146106002

The intrinsic or the extrinsic fluorescent probe attached with biomolecules displays the local and global motion because of tumbling of the biomolecules through Brownian motion. The tumbling of molecules in solution leads to depolarization of excited state. Fluorophores in the homogeneous solutions are randomly oriented in their ground state at room temperature. When a fluorophore is excited with polarized light, fluorophore has the higher probability of absorption of photons whose electric vector aligned parallel to the absorption transition dipole of the fluorophore. This process is called as photoselection. During the period of fluorophores in the excited state (~10^{-9} s), the extent of polarization changes due to the Brownian rotational motion of molecules. This extent of depolarization depends on rate of rotational diffusion and can be measured in terms of anisotropy. The fluorophore whose fluorescence lifetime is comparable to its rotational rate displays non-zero anisotropy. Fluorophores tumbling very fast shows the anisotropy values close to zero.

The anisotropy can be expressed as

$$r_{ss} = \frac{I_{\parallel} - GI_{\perp}}{I_{\parallel} + 2GI_{\perp}} \qquad 2.24$$

Here r_{ss} is steady-state anisotropy, $I_{\parallel}$ is the fluorescence emission intensity when emission and excitation polarization are vertically aligned and $I_{\perp}$ represents the fluorescence emission intensity when the emission polarizer aligns perpendicular to the excitation polarizer. The schematic of fluorescence anisotropy measurement is displayed in **Figure 2.6.** G is known as the geometry factor (G-factor) and is defined as the ratio of monochromator and detector sensitivity to intensity of vertically (S_V) and horizontally (S_H) polarized light.

$$G = \frac{S_V}{S_H} \qquad 2.25$$

The value of G factor depends on the emission wavelength and bandpass of the monochromator. Anisotropy is a dimensionless quantity and is independent of the sample concentration (in absence of any artifact like scattering).

TH-3004_146106002

2.1.2.5 Time-resolved Fluorescence intensity decay:

Time-resolved measurement can be studied by two methods such as time domain methods and frequency domain method. In the time domain, a sample is excited with the pulse of light for an ultra-short duration. While in the frequency domain sample is excited by intensity modulated light. In this thesis, we utilized the time domain method (TCSPC) for our lifetime fluorescence measurements.

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

After exciting the sample with short pulse light, time-dependent fluorescence intensity decay of the fluorophore is collected using a technique known as the Time-Correlated Single Photon Counting (TCSPC). From the slope of log *I(t)* versus *t* plot, fluorescence decay time (τ) is measured. To circumvent the effect of anisotropy or rotational diffusion on the fluorescence intensity decay of fluorophore, the fluorescence intensity decay is commonly collected at an angle of 54.7° through a polarizer. Scheme for the time-resolved fluorescence lifetime measurement is displayed in **Figure 2.7.**

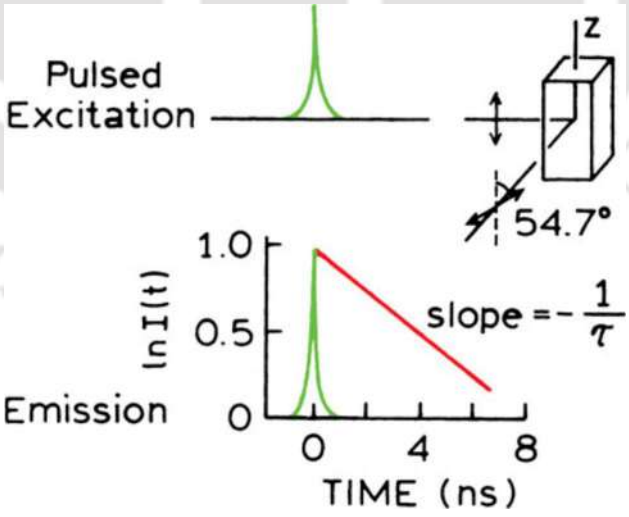


Figure 2.7 Illustration of time-domain lifetime measurement. Adapted from Principle of Fluorescence Spectroscopy[156] by J. R. Lakowicz, 3rd edition, 2006.

TH-3004_146106002

The fluorophore which has the single lifetime, intensity decay equation can be written as

$$I(t) = I_0 exp\ (-t/\tau) \qquad 2.26$$

Where, *I(t)* represent intensity at any time *t*, I_0 is the initial intensity (*t = 0*) of fluorophore and τ denote the fluorescence lifetime of the fluorophore.

The fluorescence lifetime is the average time that a fluorophore remains in the excited state prior to emission of a photon. The fluorescence lifetime for a single exponential decay can be calculated by averaging the time (*t*) over the intensity decay of the fluorophore**.**

Fluorescence lifetime *(τ)* can be expressed as follows:

$$< t >= \frac{\int_0^\infty tI(t)dt}{\int_0^\infty I(t)dt} = \frac{\int_0^\infty t\ exp(-t/\tau)dt}{\int_0^\infty exp(-t/\tau)dt} \qquad 2.27$$

The above equation can be solved to obtain the average time of a fluorophore remains in the excited state is measured to be equal to fluorescence lifetime τ.

$$< t > = \tau \qquad 2.28$$

For the fluorophore that displays multi-exponential decay, as seen for fluorophore in different environments or due to the presence of multiple fluorophores. The time-dependent intensity for a multi-exponential decay can be expressed as

$$I(t) = \sum_i \alpha_i exp\ (-t/\tau_i) \qquad 2.29$$

Here, τ_i and α_i denote the i^{th} fractional lifetime and i^{th} fractional amplitude respectively for the decay component. Addition of fractional amplitudes equals to unity and can be expressed as

$$\sum_i \alpha_i = 1 \qquad 2.30$$

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

TH-3004_146106002

$$\tau_m = \sum_{i=1}^{n} \alpha_i \tau_i \qquad 2.31$$

Mean lifetime (τ_m) is directly proportional to the area under the intensity decay curve. As steady-state measurement provides an average of the time-resolved event over the fluorescence intensity decay of the fluorophore. Knowing the relationship between steady-state and time-resolved fluorescence is also essential.

When a fluorophore shows a single decay time (τ) its intensity decay can be described as

$$I(t) = I_0 e^{-t/\tau} \qquad 2.32$$

The relationship between decay time and steady-state intensity (I_{ss}) can be expressed as

$$I_{ss} = \int_0^{\infty} I_0\, e^{-t/\tau} dt = I_0 \tau \qquad 2.33$$

In comparison with steady-state fluorescence, time-resolved measurement has many advantages. These include, A) it does not depend on the fluorophore concentration, B) tells about the microenvironment around fluorophore by revealing multiple lifetimes, and C) explains about the type of quenching like static or dynamic quenching.

2.1.2.6 Time-resolved fluorescence intensity decay analysis:

Data acquired through time-resolved fluorescence are complex and mere graphical methods cannot be sufficient to evaluate and extract the information from the data. Intensity decay curve contains three curves 1) Instrument Response Function (IRF) [$L(t_k)$], 2) Measured intensity decay $N(t_k)$, and 3) calculated (fitted) intensity decay $Nc(t_k)$. The measured intensity decay $N(t_k)$ is a convolution of the IRF [$L(t_k)$] with actual intensity decay of the sample, $Nc(t_k)$. All these functions are in terms of discrete time (t_k) since the counted photons are collected into channel each with known width (Δt) and time t_k.

The IRF, $L(t_k)$ also called as the lamp function which is an instrument response to sample with zero fluorescence lifetime. This decay indicates the shortest time that can be measured by TCSPC. The IRF can be collected by using dilute non-fluorescence scattering solution

TH-3004_146106002

like colloidal silica or $Mg(OH)_2$ in absence of any emission filter. 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 a substantial contribution from IRF. This has to be analyzed to determine the actual lifetime of the sample. The mathematical method used to extract the intensity decay data 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 expressed as

$$I_k(t) = L(t_k)I(t - t_k)\Delta t(t > t_k) \qquad 2.34$$

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

$$N(t_k) = \sum_{t=0}^{t=t_k} L(t_k)I(t - t_k)\Delta t \qquad 2.35$$

The above equation can be expressed as an integral, for small value of Δt.

$$N(t_k) = \int_0^t L(t')I(t - t')dt' \qquad 2.36$$

According to the above equation, the measured intensity at time t is given by the sum of the intensity expected for all the δ -function excitation pulses that appears until time t. The new intensity decay is generated in the sample until it is nonzero intensity in $L(t_k)$.

Numerous methods like nonlinear least square analysis (NLLS), the maximum intensity entropy method, the method-of-moments, phase-plane method and Laplace transformation etc. are used to analyze TCSPC data. The Marquardt's algorithm based non-linear least square analysis (NLLS) is widely used to analyze the complex intensity decay.

TH-3004_146106002

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, $Nc(t_k)$ 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

$$\chi^2 = \sum_{k=1}^{n} \frac{1}{\sigma_k^2}[N(t_k) - N_c(t_k)]^2 \qquad 2.37$$

$$= \sum_{k=1}^{n} \frac{[N(t_k) - N_c(t_k)]^2}{N(t_k)} \qquad 2.38$$

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 represented as

$$\chi_R^2 = \frac{\chi^2}{n-p} = \frac{\chi^2}{\nu} \qquad 2.39$$

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

If calculated and experimentally obtained data are closely matched then value of χ_R^2 will be close to unity. If the data does not fit properly, then 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.

Additional parameter to judge the quality of fit is known plots or distribution of residuals.

The difference between measured and fitted function yield 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

TH-3004_146106002

$$D_k = \frac{I(t_k) - I_c(t_k)}{\sqrt{I(t_k)}} \qquad 2.40$$

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

For MEM analysis, the intensity decay was 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 lifetime by minimizing the χ^2 and maximizing the entropy, S (known as Shannon-Jaynes Entropy) given by equation 2.41 and 2.42, respectively. The goodness of fit was determined by the value of obtained reduced chi square (χ_R^2) values as well as the increased entropy values. At least three such fits were performed for each sample and the fit with best χ_R^2 values were selected.

$$S = \sum_i -p_i log\ p_i \qquad 2.41$$

Where, p_i denotes the probability (amplitude) for the i^{th} 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 values of χ^2, MEM gives the distribution for which entropy, S is maximum. Here, χ^2 is given by

$$\chi^2 = \frac{\left(\frac{1}{2}\right)\sum_{i=1}^{m}[F_c(t_i) - F_e(t_i)]^2}{F_e(t_i)} \qquad 2.42$$

Here, *m* represents the number of data points and $F_c(t_i)$ and $F_e(t_i)$ are the calculated and experimental fluorescence, respectively at time *t*.

2.1.2.7 Förster Resonance Energy Transfer (FRET):

Resonance Energy Transfer (RET) is an important process which happens in the excited state. FRET takes place when there is an overlap between emission spectrum of a fluorophore (donor) and the absorption spectrum of a nearby chromophore (acceptor). FRET is an important method, employed for exploring the protein-protein interaction, protein folding and protein aggregation. FRET is an electromagnetic process; involving the

TH-3004_146106002

transfer of excited state energy from excited state donor (D) to ground state acceptor (A) molecules. Transfer of energy happens 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 donor and acceptor molecules, 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 is 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

$$R_0^6 = 8.79x10^{23}[k^2\eta^{-4}Q_D J(\lambda)](\text{Å}) \qquad 2.43$$

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

$$J(\lambda) = \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda M^{-1}\, cm^3 \qquad 2.44$$

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

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

$$k_T(r) = \left(\frac{1}{\tau_D}\right)\left(\frac{R_0}{r}\right)^6 \qquad 2.45$$

Where, r denotes the distance between the donor and acceptor, R_0 represents the Förster radius and τ_D is the lifetime of the donor in absence of acceptor.

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

$$E = \frac{R_0^6}{R_0^6 + r^6} \qquad 2.46$$

TH-3004_146106002

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

$$E = 1 - \left(\frac{F_{DA}}{F_D}\right) \qquad 2.47$$

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

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

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 method to perceive the conformations of protein and nucleic acids in solution. To exhibition the CD signal, chromophore should have chirality or be present in an 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 written as

$$\Delta A = A_L - A_R \qquad 2.49$$

Where A_L and A_R denotes the absorbance measured 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 the absorption spectrum by calculating the area under each CD bands and noting signals. These quantities together are known as the rotational strength used as a measure of CD intensity.

$$R_{0\alpha} = (3hc/8\pi^3 N_0) \int \{[\theta(\lambda)]/\lambda\}\, d\lambda \qquad 2.50$$

TH-3004_146106002

Where, h is Planck's constant and c denotes 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 an asymmetric molecule.

$$R_{0\alpha} = imaginary(\langle\psi_0|\tilde{\mu}|\psi_a\rangle \cdot \langle\psi_a|\tilde{m}|\psi_0\rangle) \quad 2.51$$

Here, $\tilde{\mu}$ is an electric dipoles operator and $\tilde{m}$ is magnetic dipole operator.

As CD is one of the absorption techniques, chromophores that show the CD spectrum also display absorption spectrum. When plane polarized light pass through the optically active sample, one of the circular components 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 components instead of recombined them. Relationship between circular dichroism and ellipticity can be expressed as

$$[\theta] = 3300\Delta\varepsilon \quad 2.52$$

Here, $[\theta]$ denote the molar ellipticity and most of circular dichroism measurements are based on ellipticity determinations.

The far-UV CD spectra basically divulge the information about the secondary structural content of the protein like $\alpha - helix, \beta - sheets$ and random coil. In the region of 240 to 190 nm, light is majorly absorbed by the peptide bond and exhibits a strong π *to* π^*(~190 nm) and weak but broad n to π^* transitions (210-220 nm).

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

TH-3004_146106002

2.2 Materials:

Chloramphenicol (C3175); Ampicillin (A8351); Isopropyl β-D-1-thiogalactopyranoside (I5502); Acrylamide (A3553); Agarose (A9539); EtBr (E7637); Phenylmethanesulfonyl fluoride (P7626); β-Mercaptoethanol (63689); Glycine (50046); Sinapic Acid (85429); $CaCl_2$ (C8106); Gel loading buffer (G2526); Trifluoroacetic Acid (T6508); 2,2,2-Trifluoroethanol (T8132); N-acetyl tryptophan amide (A65501); Ammonium persulfate (A3678); N,N,N',N'-Tetramethylethylendiamine (A7024); Sodium Pyruvate (P2256); Nicotinamide adenine dinucleotide (N8129); Lactate Dehydrogenase (10127876001); Bovine serum albumin (A3059) and Hen Egg-white Lysozyme (L6876) protein were 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 (Carlbad, California, USA). Luria Broth (M1245); Luria Agar (M557); Terrific Growth Medium (G004) and imidazole (GRM1864) were procured from HiMedia Laboratories, India. The PD-10 desalting column (17-0851-01) and Ni Sepharose His-Trap FF affinity columns (17525501) were purchased from GE Healthcare. The TRIS (Hydroxy methyl amino ethane) (93315); Citric acid monohydrate (100244); Sodium acetate (17952); di- Sodium hydrogen phosphate (17549); Sodium Chloride (40731); Dimethylformamide (17754); Dimethyl sulfoxide (61857); Sodium Dodecyl Sulfate (184190); Potassium dihydrogen phosphate (104873); Potassium chloride (91635) were procured from Merck India Limited. QuikChange II Site- directed Mutagenesis Kit (20024) was procured from Agilent Technologies. His trap column (17-5255-01) was purchased from GE Healthcare. All other analytical grade buffer and salt chemicals were purchased from Merck with purity of ≥98%. All chemicals used were analytical grade with purity ≥98%.

Mass Ruler DNA Ladder Mix (SM04403) from Thermo Scientific; Ultra Low Range Molecular Weight Marker (1,060-26,600) (M3546) from Sigma Aldrich Chemicals Pvt. Limited, Bengaluru, India.

TH-3004_146106002

2.3 Experimental Methods:

2.3.1 Cloning of DHN1 gene in pET19b vector:

The DHN1 gene was cloned into the pET19b expression vector between *BamHI* and the *NcoI* restriction sites[149].

2.3.2 Site-Directed Mutagenesis:

As per the designed plan for mutation, primers were designed using the Agilent QuickChange Primer designing software and the T_m values obtained were verified using the formula

$$T_m = 81.5 + 0.41(\%GC) - (675/N) - \%mismatch \qquad 2.53$$

Where N represents the primer length in bases and the values for % GC and % mismatch are whole numbers. The primer was procured from Integrated DNA Technology, Inc in lyophilized dry powder form. Subsequently, the stock solutions were made in 20 mM Tris

Table 2.2 Shows the reaction mixture composition that was used for the PCR thermal cycling.

Control reaction	Sample reactions
5 µL of 10x reaction buffer	1 µL of 10x reaction buffer
2 µL of pWhitescript 4.5-kb control plasmid (5 ng/ µl)	0.4 µL (5-50 ng) of dsDNA template
1.25 µL (125 ng) of oligonucleotide control primer #1 [34-mer (100ng/ µL)]	0.25 µL (125 ng) of oligonucleotide primer #1
1.25 µL (125 ng) of oligonucleotide control primer #2 [34-mer (100ng/ µL)]	0.25 µL (125 ng) of oligonucleotide primer #2
1 µL of dNTP mix	0.2 µL dNTP mix
38.5 µL ddH_2O (to bring the final reaction volume to 50 µL)	8.8 µL ddH_2O to a final volume of 10.9 µL

Followed by the addition of 1 µL PfuUltra HF DNA polymerase (2.5 U/ µL) to control and sample reactions.

TH-3004_146106002

Table 2.3 The parameters used for the PCR 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

buffer (pH 8.0) and stored at -20 °C. The parent DNA plasmid (DHN1) from *Zea mays* (kindly gifted by Prof. T.J. Close, University of California, USA) was used as the template for mutagenesis. For the mutation, the QuickChange II Site-directed mutagenesis Kit (20024) was used with slight modifications. The sample composition used mutation for PCR reactions is shown in **Table 2.2** and the parameters used for PCR reaction is shown in **Table 2.3**.

After the PCR reaction using a thermal cycler (Applied Biosystems), samples were placed on ice for 2 minutes to cool the reaction to ≤37 °C. For the final step of this mutagenesis process, 1μL of the *Dpn* I restriction enzyme (10 U/ μL) is added directly to each amplification using a small pipet tip. Each reaction mixture is gently and thoroughly mixed by pipetting the solution up and down several times. The sample mixture is then spun in a microcentrifuge for 1 minute and immediately incubated each reaction at 37°C for 1 hour to digest the parental (i.e., the template) supercoiled dsDNA. Subsequent to the *Dpn* I digestion, the daughter plasmids containing the desired mutation are then transformed into the DH5α bacterial strain and incubated at 37 °C for 12 hours. The bacterial colonies obtained after transformation were further screened for the desired mutation by DNA sequencing through Eurofins Pvt Ltd.

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

TH-3004_146106002

2.3.3 Competent cell preparation and transformation of plasmid DNA:

E. coli DH5α and BL-21 Codon$^+$ (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 (Chloramphenicol in case of BL-21 Codon$^+$ (DE3)). The culture was allowed to grow in a shaking incubator at 37 °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 °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 re-suspended gently by pipetting or inverting in 15 mL of ice cold 80 mM $MgCl_2$ and 20 mM $CaCl_2$ solution. Then, cells were centrifuged again at 3000 rpm for 10 minutes at 4 °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 °C until further use. All the materials used for the competent cell preparation were autoclaved and the all the steps were performed in laminar air flow to avoid any contamination.

The plasmid containing gene of interest (wild-type DHN1 and its mutants DHN1 CW1, DHN1 W3) were transformed into DH5α (for amplification) and BL-21 Codon$^+$ (DE3) (for expression and purification) competent cells. For transformation following steps were followed: the *E. coli* DH5α and BL-21 Codon$^+$ (DE3) competent cells were taken out from the -80 °C and placed in ice for 10 minutes. 1-2 µL of the plasmid containing the gene of interest was added to the aliquots of thawed competent cells and kept 30 minutes in ice. One aliquot without any plasmid addition was served as the control. After 30 minutes of incubation, cells were treated with heat shock at 42 °C for 60 seconds and immediately transferred to ice and incubated for 5 minutes. 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 6000 rpm for 10 minutes. After discarding supernatant, the collected pellet was re-suspended in 100

TH-3004_146106002

µL of fresh LB media and plated on Luria Agar plates containing Ampicillin (Amp^R +$Chloramp^R$ in case of BL21 $Codon^+$ cells) as the antibiotic for selection. The LB plates were then incubated at 37 °C for the 12 hours. A range of 10-20 positive colonies were obtained for the mutant plasmids (upon transformation in DH5α cells), while transformation in BL21 $Codon^+$ cells result in several scattered colonies. Single colony from the transformed plate was selected for the amplification and expression/purification. Transformation of plasmid into the DH5α cells and BL 21 $Codon^+$ cells were performed individually.

All plasmids were isolated from transformed DHN5α cells using the Sigma GenElute™ Plasmid Miniprep Kit (PLN 70) and stored in Tris buffer pH 8.0 at -20 °C for PCR-like amplification and DNA sequencing.

2.3.4 Expression, purification of DHN1, DHN1 CW1 and DHN1 W3:

A single colony of transformed *E. coli* BL21 $codon^+$ (DE3) cells was picked up from a LB agar plate and was used for over expression of DHN1 and its mutant variants. The culture was grown in LB media, containing 100 µg/mL Ampicillin and 25 µg/mL Chloramphenicol (at 37°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, 0.25 mM) at 37 °C for 3-3.5 hours. Induced cells were centrifuged at 5000 rpm for 15 minutes at 4 °C to collect the pellet. The collected pellets were re-suspended in ice cold lysis buffer (50 mM Sodium phosphate (pH 7.4); 300 mM NaCl;10 mM Imidazole; 3 mM, 1 mM PMSF (TCEP; 3 mM, in case of DHN1 CW1)) and kept at 4 °C for 20 minutes. Subsequently, resuspended cells were disrupted by sonication at 33% power output with pulse cycle of 4 second on and 10 seconds off for 20 minutes (depends on culture volume) at 4 °C. The lysed cells were centrifuged at 12000 rpm for 30 minutes and subsequently the collected supernatant was heated at 95 °C for 30 minutes. Subsequently, the heated supernatant was ultracentrifuged (Beckman Coulter, optima TLX) using 70 Ti rotor at 117,000xg for 30 minutes at 4 °C. The collected supernatant was filtered with a 0.45 µm filter (Milipore). The filtered supernatant was then allowed to bind to a pre-equilibrated (binding buffer 10 mM imidazole, 50 mM Sodium phosphate (pH 7.4); 300 mM, 3mm B-ME) His Trap FF SP column for 4-5 hours at room

TH-3004_146106002

temperature (298 K). After binding, column was washed with wash buffer (50 mM Sodium phosphate (pH 7.4); 300 mM NaCl; 20 mM Imidazole) followed by elution with elution buffer (50 mM Sodium phosphate (pH 7.4); 300 mM NaCl and 250mM Imidazole).

Purity of proteins was ascertained by SDS polyacrylamide gel electrophoresis (15% Acrylamide). The fractions which showed protein band in SDS-PAGE were pooled and dialyzed against water in case of DHN1 and buffer (50 mM Sodium Phosphate,50 mM NaCl, pH 7.4) in case of DHN1 mutants. Proteins were further analyzed with mass spectrometry to ascertain the identity and purity of proteins.

2.3.5 Sequence-based analysis:

2.3.5.1 Amino acid composition and disorder prediction plot:

Sequence based analysis of DHN1 and its mutants DHN1 CW1 and DHN1 W3 were performed using the Composition profiler tool server (*http://www.cprofiler.org/index.html*). The SwissPort version 51 protein database[159] was used as a background data set to compare amino acid composition of DHN1 and its mutants DHN1 CW1 and DHN1 W3. Intrinsic disorder prediction plot for DHN1 and its mutants DHN1 CW1, DHN1 W3 were executed using Genesilico MetaDisorder server[160] (*http://iimcb.genesilico.pl/metadisorder/index.htm)*. Gensilico metaDisorder server predicts the disorderness in proteins by using the different available predictors, like IUPred, RONN, Pdisorder, MetaDisorderMD and VSL2.

2.3.5.2 Charge Hydropathy plot and hydrophobic cluster analysis:

Charge hydropathy plots (CH plot) for the DHN1 and its mutants DHN1 CW1, DHN1 W3 were drawn using the PONDR server. The value of hydrophobicity <H> and the mean net charge <R> was calculated using the PONDR server as described by the Uversky and co-workers[54] in 2000. The charge hydropathy plot is divided in two regions such as intrinsically disorder protein region and ordered protein region, respectivelyby a line. Equation used to divide CH plot by a line can be expressed as $< H >_{boundary} = (< R > +1.151)/2.785$. Hydrophobic cluster analysis (HCA) of DHN1 and its mutants DHN1 CW1 and DHN1 W3 were performed using the program DRAWHCA 1.0.2 as described by Callebaut and co-workers[161] in 1997 (*http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::HCA*).

TH-3004_146106002

2.3.6 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE):

The reducing SDS-PAGE was used as per the protocol described in Sambrook and Russel, Molecular Cloning-A laboratory manual[162] to examine the purity of purified protein. Resolving gel (15% acrylamide) was utilized for running all the purified protein samples. The composition of SDS-PAGE is given in appendix. Protein samples were prepared for loading by mixing with the gel loading buffer in 4:1 ratio and then allowed to boil at 95 °C for 5 minutes. Subsequently, samples were separated on vertical Mini-PROTEAN^{R} Tetra Electrophoresis System (Bio-Rad) at 80 Volts for 2.15 hours. For visualization of separated sample, gels were removed from electrophoresis apparatus and stained with colloidal Coomassie stain for 4 hours and then the gels were placed in de- staining solution (**see appendix**) overnight to allow the de-staining of gel. Standard molecular markers were used to determine the approximate molecular weight of protein in the gel.

2.3.7 Protein Estimation:

For the determination of protein concentration, the Lowry method was used [163]. The standard curve was made using BSA protein. Various dilutions of BSA ranging from 0.05 mg/mL to 1 mg/mL were prepared from freshly prepared stock of BSA (1 mg/mL) in deionized water. 200 µL of DHN1, DHN1 CW1, DHN1 W3 and each dilution of BSA, were added to the aluminum foil wrapped test tubes. All the samples were taken in duplicates. Deionized water was used as blank. Afterward 2 mL of Reagent I (**see Appendix**) was added to each of the tubes and vortexed thoroughly. After proper vortexing, all tubes were incubated for 10 minutes at room temperature. After incubation, 200 µL of Reagent II (**see Appendix**) was added to all tubes and mixed properly. Now all the tubes were incubated in the dark for 30 minutes at room temperature. Then, absorbance of each sample was measured at 650 nm to determine the protein concentration. The approximate concentration DHN1, DHN1 CW1 and DHN1 W3 were calculated from the slope of BSA concentration versus absorbance plot.

TH-3004_146106002

2.3.8 Mass Spectrometry:

Molecular mass of protein was determined by using MALDI-TOF mass spectrometer (Daltonics Bruker, Germany). For the preparation of matrix, saturated solution of Sinapic acid in TA- 30 solvent was used. TA- 30 solvent was prepared by mixing 0.1% Trifluoroacetic acid and acetonitrile in 7:3 ratio. Further, Sinapic acid was allowed to sonicate for 30 minutes and subsequently centrifuged at 12000 rpm at room temperature (298 K) for 15 minutes. Supernatant was then mixed with protein samples in 2:1 ratio and spotted on the target plate and then allowed to dry at room temperature for 2 hours. Data acquisitions of all sample were done in linear mode with flexControl and analysis was done with flexAnalysis software from Bruker Daltonics.

2.3.9 UV-Visible Spectra:

The absorption spectra of DHN1, DHN1 CW1 and DHN1 W3 were recorded using double beam Lambda-25 UV-Vis Spectrophotometer (Perkin Elmer, USA) at room temperature (25 °C). UV transmissible quartz cuvette of 10 mm path length (Hellma; Z600210) was used for all absorption measurements. The absorption spectra were acquired by scanning (multiple scan; 3−5) sample between 250 and 800 nm with 1 nm resolution and a scan speed of 480 nm/minute and averaged subsequently. All the spectra displayed are averages of at least two independent measurements. To determine the molar extinction coefficients, DHN1, DHN1 CW1 and DHN1 W3 at different dilutions were dissolved in deionized water. DHN1 CW1 sample always contained 5 mM TCEP in all absorption measurements. Blanks were identical to the sample in all respects with the exception that they contained no protein.

For recording pH dependent absorption spectra, DHN1 protein was dissolved in different buffers (25 mM) as described under circular dichroism measurements. To monitor the effect of salts on absorption spectra, DHN1 protein was dissolved in 250 mM KCl or NaCl.

To probe the temperature dependence of absorption spectra, DHN1 protein was dissolved in deionized water. Temperature of protein samples were maintained using a circulating water bath (Julabo; FPW50-HL). The cuvette containing the sample was sealed with parafilm and incubated in water bath for at least 20 minutes at chosen temperature prior to

TH-3004_146106002

collecting spectra. Initially absorption spectra were measured at 25 °C. Then,absorption spectra were collected by heating the sample to 85 °C. As the sample cooled, absorption spectra were recorded at 65 °C, 45 °C and 25 °C, respectively. All spectral acquisition parameters were kept identical for each temperature.

2.3.10 Labelling with external probe:

Labeling of single cysteine in the DHN1 CW1 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[164]. 60 mM stock of IAEDANS was prepared by solubilizing the probe into the DMF. 300 µM of DHN1 CW1 was dissolved in 1 mL labeling reaction buffer (25 mM Tris.HCl (pH 9.3); 25 mM NaCl; 5 mM TCEP). 50 µL of 60 mM IAEDANS in DMF was added (final concentration of IAEDANS was 3 mM) to 1 mL of 300 µM DHN1 CW1. 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 25 mM Tris.HCl (pH 9); 25 mM NaCl. 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 is 6100 $M^{-1}cm^{-1}$ at 337 nm[165]. Dye/protein ratio of 0.9 was obtained for DHN1 CW1 mutant protein.

2.3.11 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 acquiring the luminescence emission spectra for DHN1 protein, the protein was excited at different wavelengths: 310, 340, 355,375 and 410 nm with slit width of 2 nm for excitation and 15 nm for emission. Emission spectra were recorded between the 330-600, 360-650, 375-650, 390-650, and 430-750 nm for excitation wavelengths 310, 340, 355, 370 and 410 nm, respectively. Then, all the spectra were normalized to unity and plotted for comparison. Excitation spectra were recorded for DHN1 protein at three different wavelength 425, 450 and 480 nm for excitation wavelength between 320-410, 430 and 460

TH-3004_146106002

nm, respectively. The excitation spectra were acquired with excitation slit width of 2 nm and emission slit width of 5 nm.

For all tryptophan steady state fluorescence intensity measurements of DHN1 CW1 and DHN1 W3, samples were excited at λ_{ex} 295 nm (slit width = 1 nm) and emission was collected between 310 nm to 500 nm (slit width = 5 nm). For acquiring tryptophan steady state anisotropy of DHN1 CW1 and DHN1 W3; all samples were excited at fixed 295 nm (slit width = 1 nm) and emission was collected at fixed 346 nm (slit width = 5 nm).

For acquiring steady state fluorescence intensity of dansyl labeled with cysteine of DHN1 CW1, all samples were excited at λ_{ex} 340 nm (slit width = 1 nm) and emission was collected between 355 nm to 670 nm (slit width = 5 nm). For collecting steady state anisotropy of dansyl labeled with cysteine of DHN1 CW1, all samples were excited at fixed λ_{ex} 340 nm (slit width = 1 nm) and emission was collected at fixed λ_{em} 503 nm (slit width = 5nm).

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 °C) in quartz cuvette (Hellma; Z802875) with a 10 mm pathlength. The monochromator calibration was verified by cross checking the emission maxima of standard samples (N-acetyl tryptophan amide (NATA) at λ_{ex} 295 nm; Pyrene at λ_{ex} 310 nm; and Fluorescein at λ_{ex} 492 nm) with reported values.

To study the SDS induced structural transitions of DHN1 protein, 20 µM of DHN1 CW1 and DHN1 W3 were dissolved water were mixed in different SDS concentration (0.1 -20 mM). N-acetyl tryptophan amide (NATA) was also employed as a control under identical conditions for all experiments.

2.3.12 Stokes shift measurements:

Excitation wavelength dependent change in Stokes shift (in wavenumber units) was computed from the emission spectra of the DHN1 protein when excited at 310, 340, 355,

TH-3004_146106002

370 and 410 nm using the equation 2.20. Here,we used wavenumber (cm^{-1}) of λ_{ex} instead of λ_{abs}.

2.3.13 Quantum yield measurements:

Quantum yield for the DHN1 protein was calculated by excitation at 355 nm using the 9, 10 diphenyl anthracene (DPA) as reference[166], The quantum yield was calculated using following equation

$$\Phi_f^i = \frac{F^i f_s n_i^2}{F^s f_i n_s^2} \Phi_f^s \qquad 2.54$$

Where, Φ_f^i and Φ_f^s denotes the photoluminescence quantum yield of the sample and that of the standard, respectively. $F^i \; and \; F^s$ are the integrated luminescence intensities of sample and standard spectra, respectively. f_i denotes the absorption factor (absorbance), the fraction of light impinging on the sample that is absorbed ($f_x = 1 - 10^{-A_x}$, where A = Absorbance). $n_s \; and \; n_i$ represents the refractive indices of the reference and the sample.

The excitation at 355 nm was selected because it is close to maximal emission yield for DHN1 protein and far red-shifted from the intrinsic excitation bands of aromatic amino acids. Absorbance spectra was recorded for DHN1 protein in deionized water whereas for the 9,10-DPA (diphenylanthracene), measurement was done in cyclohexane (99%). The concentration of sample was adjusted to match the absorbance value of reference (DPA) at 355 nm (0.01-0.03). Similarly, the luminescence spectra of DHN1 protein dissolved in deionized water and DPA in cyclohexane (99%) were measured. Excitation of the sample and reference were done at 355 nm with slit width of 1 nm and nearly full emission spectrum was recorded between 370-690 nm with a slit width of 5 nm.

Quantum yield calculation was corrected for difference in refractive index of cyclohexane (DPA) and aqueous sample. For calculation of quantum yield the integrated luminescence (area under the curve) was calculated between 370 -690 nm. All the measurements were done at 25 °C in triplicates and then averaged.

TH-3004_146106002

2.3.14 Time-resolved fluorescence measurement:

All time resolved luminescence 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. The excitation of sample was done using excitation pulse light source A) 295 nm DeltaDiode^{TM} having about 0.81 ns Instrument response function (IRF) FWHM and a 20MHz repetition rate. B) 340 nm DeltaDiode^{TM} having about 0.80 ns IRF FWHM and a repetition rate of 20MHz. IRF was measured using scattering solution of colloidal silica, without emission filter for each of the sample.

To measure the ProCharTS luminescence intensity decay of DHN1 protein. The samples were excited using excitation source either of 295 nm or 340 nm. The total peak counts were collected up to 15,000 counts and luminescence intensity decay was collected in 4096 channels with temporal resolution of 0.028 ns/ channel. The signal acquisition was done with a Photo Multiplier Tube (PPD-900; Horiba Scientific) subsequent to 320 nm (Schott; WG320) or 370 nm (KV30) long pass filter for 295 nm and 340 nm excitation source, respectively. To measure the luminescence all the sample were freshly prepared in deionized water or respective buffers.

To measure the fluorescence intensity decay of DHN1 CW1 and DHN1 W3 tryptophan, samples were excited with pulsed light source 295 nm. The total peak counts were collected up to 20,000 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 decay of dansyl, labeled with cysteine of DHN1 CW1, samples were excited with pulsed light source 340 nm. Emission of dansyl were collected with 370 nm longpass filter 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.

The absorbance of all sample was kept ≤0.08 at the given excitation wavelengths to avoid inner filter effect. Emission of all samples were collected at magic angle position of 54.7°

TH-3004_146106002

with emission polarizer. All lifetime measurements were done at room temperature (25 °C) and at least three independent measurements were recorded.

Lifetime of samples were extracted by analyzing the intensity decay curve using both the discrete exponential analysis method[167] and Maximum Entropy Method (MEM)[168,169]. The software was kindly provided by Prof. N. Periasamy, Tata Institute of Fundamental Research (TIFR), Mumbai, India. The goodness of fit was evaluated by analyzing the reduced χ_R^2 and randomness of residuals.

2.3.15 FRET measurement:

Intramolecular distance in DHN1 CW1, between the tryptophan (present at 122th position) and cysteine (present at 62th position) labeled with dansyl was measured at different concentration of SDS (0.1 – 20 mM). To calculate the distance between tryptophan and cysteine at different concentration of SDS, fluorescence lifetime of tryptophan, in absence of acceptor (τ_D) and in presence of acceptor (τ_{DA}) was monitored, respectively. The fluorescence lifetime of tryptophan was collected by TCSPC. The tryptophan was excited with pulsed light source, 295 nm DeltaDiode^{TM} and then emission of tryptophan was collected with 330-355 nm bandpass filter (ET340/40X) supplied by *CHROMA,* to avoid the emission light coming from the dansyl. Total peak counts were measured up to 20,000 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 (298 K) and for each sample at least three independent lifetime measurements were done.

For acquiring SDS dependent lifetime of DHN1 CW1 tryptophan with and without acceptor, 20 μM of unlabelled/labeled DHN1 CW1 was mixed in different concentration of SDS (0.1 – 20 mM). Similarly, for measuring the SDS dependent lifetime of DHN1 W3 tryptophan, 20 μM DHN1 W3 protein was mixed in different concentration of SDS (0.1 – 20 mM).

TH-3004_146106002

Förster distance (R_0) 22 Å, for tryptophan-IAEDANS pair was used[170]. FRET efficiency (E) was calculated from mean lifetime of donor alone (τ_D) and in presenceof acceptor (τ_{DA}) by using equation 2.48.

2.3.16 Circular Dichroism analysis:

Far-UV CD spectra of DHN1 and its mutant DHN1 CW1 and DHN1 W3 were recorded on a CD spectrometer (Jasco, Model J-1500, Jasco Inc., Maryland, USA), equipped with a Peltier thermoregulator. CD spectra were recorded for all the proteins at room temperature (298 K) using 7 µM of DHN1 (dissolved in water), DHN1 CW1 and DHN1 W3 (both dissolved in 10 mM Sodium phosphate buffer; pH 7.4) protein concentration.

To monitor the effect of 2,2,2-trifluoroethanol (TFE) on structure of DHN1, CD spectra was recorded. DHN1 and its mutants DHN1 CW1 and DHN1 W3 proteins (dissolved in 5 mM Sodium phosphate (pH 7.4); 1 mM TCEP (in case of DHN1 CW1)) were mixed and incubated at least for 5 minutes with the increasing concentration of TFE (5-90% (v/v) to the final concentration) prior to collecting the CD spectra.

To measure CD spectra at different concentration of SDS, DHN1 and its mutant DHN1 CW1 and DHN1 W3 (dissolved in water) were mixed with the different concentration of SDS (0.1—20 mM).

For measuring CD spectra at various temperatures (25—90 °C), DHN1 and its mutant DHN1 CW1 and DHN1 W3 were dissolved in deionized water (containing 1 mM TCEP, in case of DHN1 CW1). All samples were incubated at least for 2 minutes at each temperature prior to collecting the CD spectra. 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. The analysis of collected CD spectra from 190 to 260 nm were done by using DichroWeb server. CDSSTR

TH-3004_146106002

program with reference set 7 was used to estimate the secondary structure content of proteins. The experimental curves were well superimposed on reconstructed curves[171–174].

2.3.17 Cryoprotection Assay:

The LDH activity assay was performed as described by earlier with slight modification[175–177]. Ten freeze/thaw cycles (15 seconds in liquid N_2 and 5 minutes at 25°C in water bath) of LDH were performed with and without DHN1 in different molar ratio (LDH: DHN1) i.e.,1:1(3 µM:3 µM), 1:2 (3 µM:6 µM), 1:4 (3 µM:12 µM) and 1:10 (3 µM:30 µM). For monitoring the oxidation of NADH at $\lambda_{340\ nm}$, LDH (0.2 µg/ml, final concentration) from treated and untreated sample were mixed with 5 mM of sodium pyruvate and 0.1 mM of NADH in 10 mM sodium phosphate buffer pH 7.4 and immediately decrease in $A_{340\ nm}$ was recorded for 300 seconds at 25 °C with fixed delay time of 30 seconds. LDH activity was determined by measuring the initial rate of decrease (slope of first 30 seconds) in the absorbance of NADH. The activity of unfrozen/untreated LDH enzyme was considered as 100% activity. All the activity in presence of DHN1, HEWL (Hen Egg White Lysozyme) and BSA (Bovine Serum Albumin) are shown with respect to activity of unfrozen/untreated LDH enzyme

2.3.18 Heat-protection Assay:

The heat-protection activity performed using the protocol previously reported[52] with some modifications. LDH was incubated with and without DHN1, HEWL, and BSA at 45 °C and 50 °C in water bath for 70 minutes in different molar ratio (LDH: DHN1) i.e., 1:1 (3 µM:3 µM), 1:2 (3 µM:6 µM), 1:4 (3 µM:12 µM), 1:10 (3 µM:30 µM) and 1: 20 (3 µM:60 µM). For monitoring the oxidation of NADH at $\lambda_{340\ nm}$, LDH (0.2 µg/ml, final concentration) from treated and untreated sample were mixed with 5 mM of sodium pyruvate and 0.1 mM of NADH in 10 mM sodium phosphate buffer pH 7.4 and immediately decrease in $A_{340\ nm}$ was recorded for 300 seconds at 25 °C with fixed delay time of 30 seconds. LDH activity was determined by measuring the initial rate of decrease (slope of first 30 seconds) in the absorbance of NADH. The activity of unheated/untreated LDH was considered as 100% activity. All the activity in presence of DHN1, HEWL, and BSA are shown with respect to activity of untreated LDH enzyme.

TH-3004_146106002

Chapter 3

Generation of DHN1 tryptophan – cysteine mutant, protein expression, purification and characterization

TH-3004_146106002

3.1 Introduction:

The selection of correct technique is an essential criterion for deciphering the structural and functional characteristics of proteins. Intrinsically Disordered Proteins (IDPs) / Intrinsically Disordered Regions (IDRs) are devoid of intrinsic spectral probes like tryptophan and richness of charged amino acid restricts their structural analysis to very few techniques like circular dichroism and NMR. Thus, it needs minor modifications like amino acid substitution, deletion, insertion and labelling with an external probe. Fluorescence spectroscopy is well-known for studying protein with an intrinsic fluorophore, i.e., aromatic amino acids [125,156,167,178–180] and external fluorophores[181,182]. The most explored and well-exploited intrinsic fluorophore in proteins is tryptophan (Trp)[156,183]. Similarly, cysteine (Cys) is frequently used to covalently link the external fluorescent probe i.e., Dansyl iodoacetamide derivative[156,184,185,186]. To elucidate the structural dynamics of protein using Förster Resonance Energy Transfer (FRET) measurement, the single Trp − Cys (labelled with external fluorescent probe) pair is widely used[156].

In-vitro mutagenesis is a vital technique that plays an essential role in characterizing the dynamics and complex relationship among the protein structure-function (discussed in chapter 2). Wild type DHN1 protein is devoid of Trp and Cys residues, which is essential for FRET distance measurement, and structural transitions. Hence, we used the in vitro site-directed mutagenesis (SDM) to incorporate the Trp and Cys in DHN1 protein.

Single tryptophan- cysteine mutant was generated to investigate the structural dynamics of DHN1 using FRET studies. Details of FRET technique is described in chapter 2. In this chapter, we used the in-vitro mutagenesis to generate the single Trp - Cys and single Trp mutants of DHN1 protein. Subsequently, the DHN1 and its mutants were expressed and purified to understand its structure and disorder characteristics. The detailed studies of the structure and disorder behavior of DHN1 and its mutants (described in next section) were carried out using computational and biophysical techniques. Both computational and experimental data indicate that the DHN1 and its mutants are fully disordered. Further, a single and double mutation in DHN1 protein does not cause any structural perturbation.

TH-3004_146106002

3.2 Selection of sites for mutagenesis

To investigate the structural and functional role of DHN1 protein by fluorescence-based methods, a prerequisite is to have a suitable fluorophore in the protein. DHN1 protein is devoid of tryptophan and cysteine residues, a widely used intrinsic fluorophore and used to label the external fluorophore, respectively. Hence, it was decided to incorporate the single Trp and Cys residue in DHN1 protein sequence via site-directed mutagenesis. Prior to choosing the site for incorporation of desired amino acids residue following considerations were taken into account. First, the mutation site must be selected in such a way that it does not perturb the conserved segments (YSK_2) in DHN1 protein, which is known to have an essential role in the function of protein such as cryoprotectant in stressful conditions [187]. Second, the appropriate distances between the donor (Trp) and acceptor (dansyl) FRET pair, which determines the FRET efficiency between them. For Trp and Dansyl (used in all FRET experiments), the reported R_0 value for FRET is 22 Å [188]. Hence the mutation site selected was done on the basis of the above-mentioned consideration.

As discussed in chapter one, it was significant to deduce the local transitions of proteins in presence of ligand binding. To investigate this, tryptophan as an intrinsic fluorescence donor and an acceptor probe (dansyl) labelled to a single cysteine residue was utilized in our studies to obtain sensitive and reproducible information through FRET-based experiment.

The above-stated procedure seemed most appropriate to elucidate the structural dynamics of DHN1 and which is only possible after the substitution of amino acid with tryptophan and cysteine at desired sites. Hence, the DHN1 protein sequence was scanned for sites suitable to be substituted by cysteine and tryptophan amino acid.

Tryptophan was substituted with tyrosine near to K-Segment, which is subsequently suitable for understanding the local and global motion upon binding with ligands [108,110].

Previous literature suggests that the substitution of alanine with cysteine is a tolerable mutation in several proteins due to the lack of any charged or bulky side chain in alanine [189]. Cysteine in the polypeptide chain inclines to behave slightly hydrophobic, unlike the free polar residue[190]. Hence replacement of small and hydrophobic alanine residue does not

TH-3004_146106002

cause any extreme structural and functional perturbation in protein. Furthermore, the relative abundance, distribution and appropriate distance from tyrosine (which further substituted by tryptophan) residue make alanine the most suitable amino acid for substitution. Therefore, alanine (A62) was selected to replace with the single cysteine residue in DHN1. While planning the mutation site, it was also ascertained that the mutation site must be away from the conserved segments of DHN1 to avoid any probable perturbation in the structural or functional activity of the protein.

Tyrosine (122) was replaced by tryptophan to generate the single tryptophan variant of DHN1. Subsequently, the alanine (62) was replaced by cysteine in the single tryptophan variant of DHN1 protein. Similarly, tyrosine (3) was replaced by tryptophan to generate the single tryptophan (at the N-terminal of protein) variant of DHN1. Both the mutants, with single cysteine (Cys 62) – single tryptophan (Trp122), and single tryptophan (Trp 3) were named DHN1 CW1 and DHN1 W3, respectively. Both mutant DHN1 CW1 and DHN1 W3 were sequenced, expressed and purified successfully as discussed in the following sections.

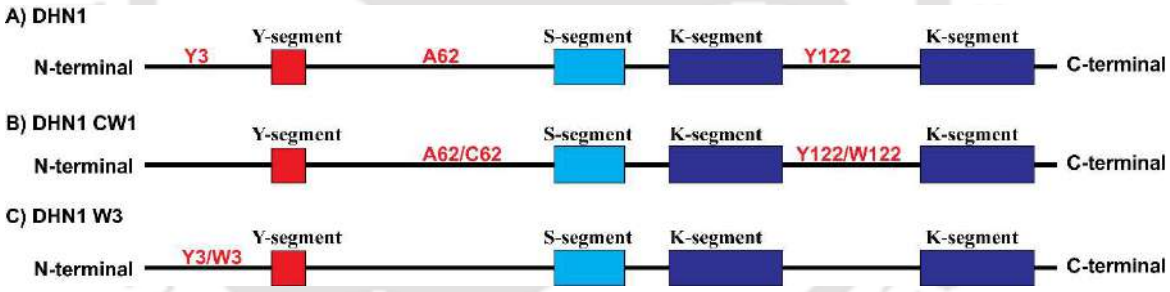


Figure 3.1. Scheme of DHN1 mutation: Substitution mutation was done by substituting the A/C and Y/W in DHN1.

TH-3004_146106002

3.3 Results and Discussion:

Sequences of DHN1 and its mutant protein

a. DHN1 [167] (16955 Da)

MEYGQQGQHG HGATGHVDQY GNPVGGVEHG TGGMRHGTGT GGMGQLGEHG GAGMGGGQFQ PAREEHKTGG
ILHRSGSSSS SSSEDDGMGG RRKKGIKEKI KEKLPGGHKD DQHATATTGG AYGQQGHTGS AYGQQGHTGG
AYATGTEGTG EKKGIMDKIK EKLPGQH

b. DHN1 CW1 [167] (16965 Da)

MEYGQQGQHG HGATGHVDQY GNPVGGVEHG TGGMRHGTGT GGMGQLGEHG GAGMGGGQFQ PCREEHKTGG
ILHRSGSSSS SSSEDDGMGG RRKKGIKEKI KEKLPGGHKD DQHATATTGG AWGQQGHTGS AYGQQGHTGG
AYATGTEGTG EKKGIMDKIK EKLPGQH

c. DHN1 W3 [167] (16978 Da)

MEWGQQGQHG HGATGHVDQY GNPVGGVEHG TGGMRHGTGT GGMGQLGEHG GAGMGGGQFQ PAREEHKTGG
ILHRSGSSSS SSSEDDGMGG RRKKGIKEKI KEKLPGGHKD DQHATATTGG AYGQQGHTGS AYGQQGHTGG
AYATGTEGTG EKKGIMDKIK EKLPGQH

Figure 3.2. Details of DHN1 mutation: Amino acid sequence of **a**) DHN1, **b**) DHN1 CW1 and **c**) DHN1 W3 (deep blue color represents the charged amino acid residues and red bold represents mutation).

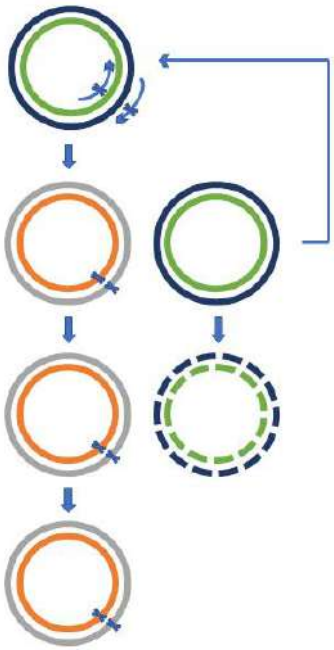
3.4 Generation of DHN1 mutants

As per the designed plan for mutation, primers were designed using the Agilent QuickChange primer design software available online. The primers generated were validated for errors with a standard formula for primer design, and T_m values obtained were validated using the formula mentioned below:

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

Where N stands for primer length (in bases) and values of % GC and % mismatch are whole numbers. The primer was procured from Integrated DNA Technology (IDT) in the lyophilized powder form. The stock solution was prepared in 20 mM Tris buffer (pH 8.0) and stored at -20°C. The basic outline followed for sit-directed mutagenesis is shown in **Figure 3.3.**

TH-3004_146106002



1. Mutant strand synthesis

Perform thermal cycling to:

- Denature DNA template
- Anneal mutagenic primers
- Extend and incorporate primers with *pfuUltra* DNA polymerase

2. *Dpn* 1 digestion of Template

Digest parental methylated and hemi methylated DNA

3.Transformation

Transform mutated molecule into competent cells

Figure 3.3. The scheme used for generation of mutation at a single site for amino acid replacement through Site-Directed mutagenesis. Parental plasmid DNA is denoted by dark blue and green spherical circles. Primers and the daughter DNA plasmid are represented as gray and orange spherical circles. An arrowhead represents the primers, and the mutation site marked as a small cross in the strands/circles. The dashed circle indicates the digested parent DNA.

Methods and the required conditions used for the PCR like reaction for the generation of mutants is described in chapter 2. Following the PCR-like amplification, the mutants plasmid was transformed into DH5α bacterial cells, as shown in **Figure 3.3**. Once the positive bacterial colonies were obtained in the LB agar (AMP^R) plates, it was subsequently confirmed by the DNA sequencing to ascertain the mutation at the desired site and to further check for unwanted modification. **Figure 3.4** shows the digested wildtype DHN1 and mutants plasmid. The generated mutants plasmid was screened based on the similar size to the wildtype DHN1 plasmid and then sent for DNA sequencing. T7 forward and reverse primers were used for DNA sequencing. The obtained DNA sequence (FASTA format) was converted into the corresponding amino acid sequence using the online available ExPASy translate tool (DNA translation tool). This process was followed for all the desired mutations.

TH-3004_146106002

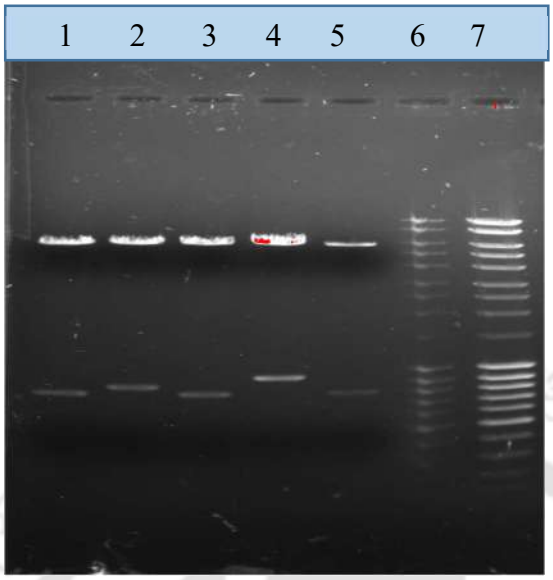


Figure 3.4. Agarose gel shows digested plasmid DNA of DHN1 from positive colonies **(lane 1- wildtype DHN1 and 2,3,4,5- mutants)** of mutant, and lane **6,7** shows DNA ladder.

3.5 Expression and purification of DHN1 and its mutants

DHN1 and its mutants DHN1 CW1, DHN1 W3 were expressed in BL 21 Codon^{+} (DE3) bacterial cells using the 0.25 mM IPTG (Isopropyl β-D-1-Thiogalctopyranoside) for induction of protein expression. **Figure 3.5A, B and C** displayed the expression of DHN1, DHN1 CW1 and DHN1 W3, respectively.

DHN1 and its mutant DHN1 CW1, DHN1 W3 were purified successfully using the boiling lysis method followed by the Ni-NTA affinity chromatography. The DHN1 and its mutant DHN1 CW1 and DHN1 W3 were allowed to bind with the Ni-NTA affinity column (5 mL). Interestingly, DHN1 and its mutants have no cloned/inserted His-Tag (a prerequisite for the Ni-NTA chromatography). DHN1, DHN1 CW1 and DHN W3 show significant binding to the Ni-NTA column. This strong binding could be due to the richness of histidine (13) amino acid residues in DHN1 and its mutants. Even washing with 50 – 100 mM imidazole

TH-3004_146106002

does not affect the binding. For elution of DHN1 and its mutant, 200 mM imidazole was required.

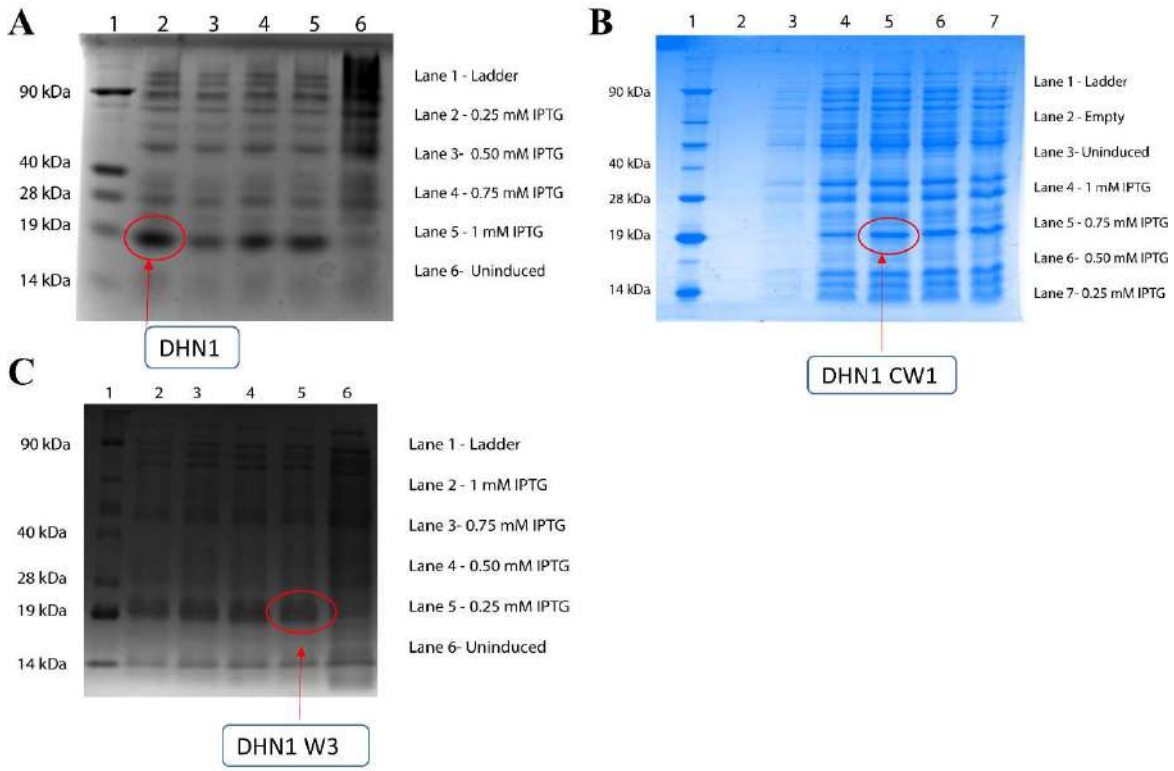


Figure 3.5. 15% SDS PAGE showing expression of (**A**) DHN1, (**B**) DHN1 CW1 and (**C**) DHN1 W3 at 37°C temperature in *E. coli* BL21 Codon^{+} strain at given concentration of Isopropyl β-D-1-thiogalactopyranoside (IPTG). Red circled band showing comparative more expression of DHN1, DHN1 CW1 and DHN1 W3 at concentration of 0.25 mM of IPTG.

Further, the purity of DHN1, DHN1 CW1, and DHN1 W3 were ascertained by reducing SDS polyacrylamide gel electrophoresis (15% acrylamide). **Figure 3.6 A, B, and C** shows the single band of purified DHN1, DHN1 CW1 and DHN1 W3, respectively on reducing SDS -PAGE. DHN1 and its mutants run on SDS-PAGE with an apparent molecular weight of approximately 19 kDa. The final concentration obtained for DHN1, DHN1 CW1 and DHN1 W3 were 2-3 mg/L of the media.

Purified DHN1, DHN1 CW1, and DHN1 W3 were further analyzed by using mass spectrometry to ascertain the identity and purity of purified proteins. The molecular mass

TH-3004_146106002

of purified DHN1, DHN1 CW1, and DHN1 W3 was measured using mass spectrometry (MALDI-TOF). **Figure 3.7 A, B and C** depicts the mass spectra of DHN1, DHN1 CW1, and DHN1 W3, respectively. **Table 3.1** shows the comparison of DHN1, DHN1 CW1 and DHN1 W3 actual Molecular Weight (MW) determined by mass spectrometry with their calculated MW. The exact MW obtained from mass spectrometry of DHN1, DHN1 CW1 and DHN1 W3 were 16951.45 Da, 16965.29 Da, and 16977.01 Da, respectively. The actual MW of DHN1 and its mutants is nearly same as MW calculated from the amino acid sequences of all the proteins.

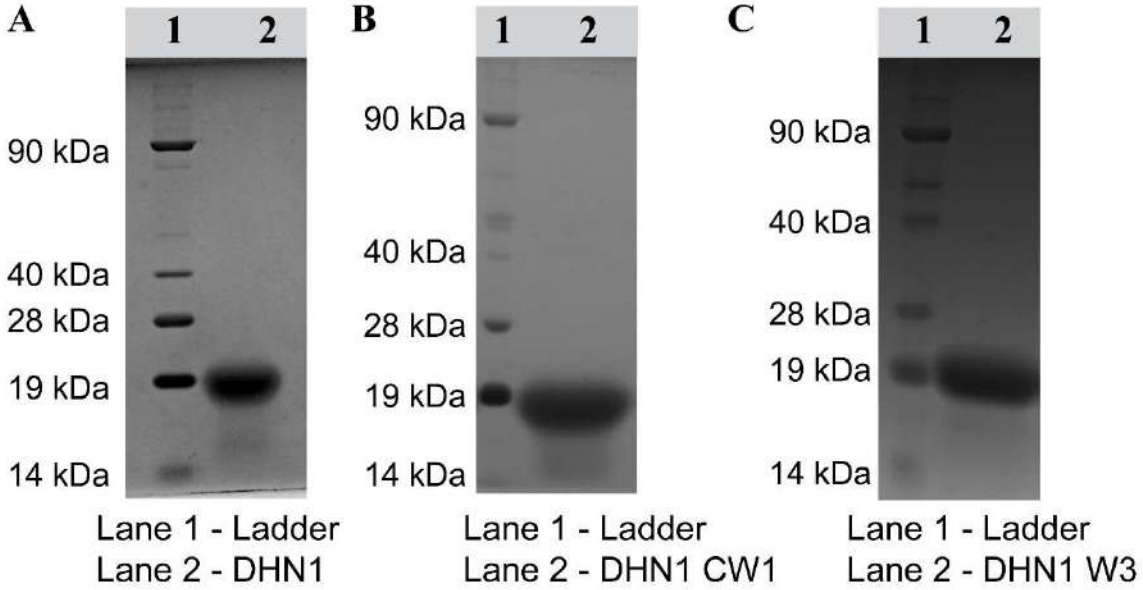


Figure 3.6. 15% Reducing SDS PAGE showing a single band of purified proteins (A) DHN1, **(B)** DHN1 CW1, and **(B)** DHN1 W3.

TH-3004_146106002

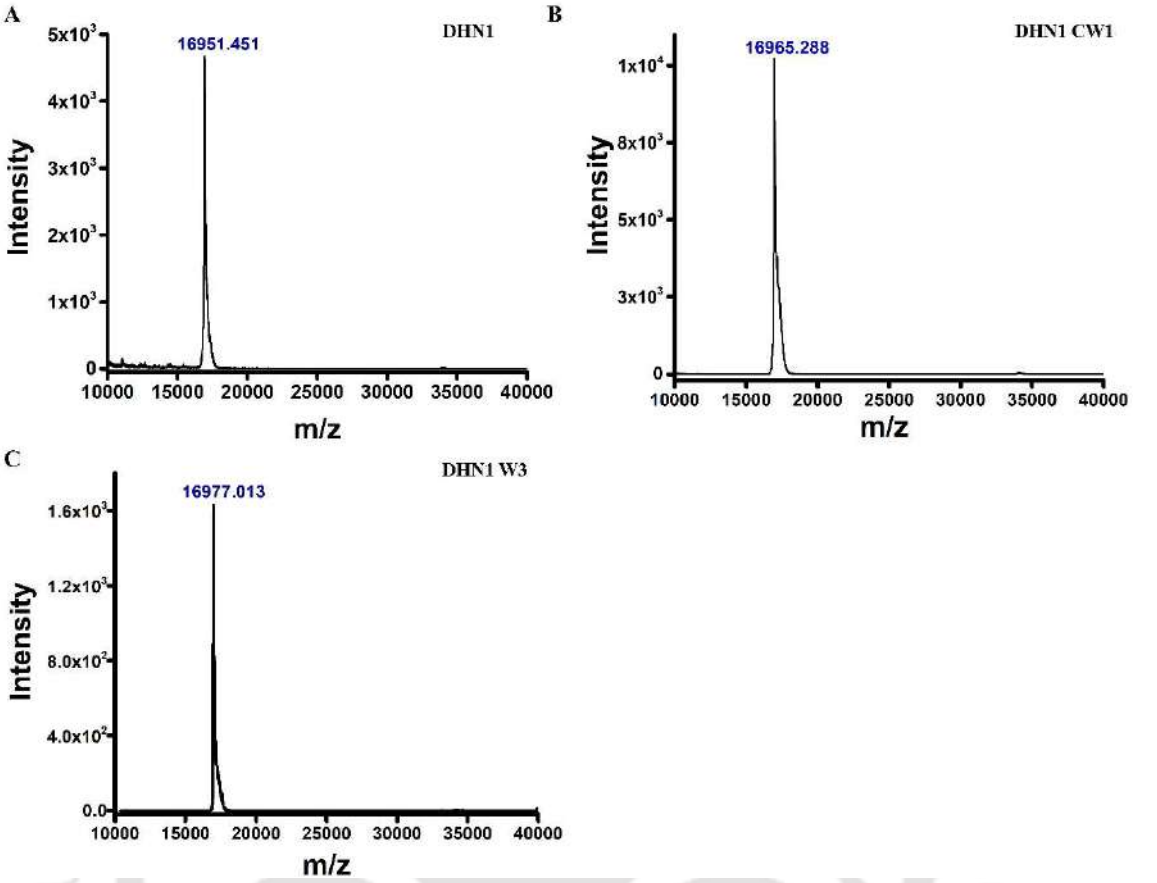


Figure 3.7 Mass of spectra of **(A)** DHN1, **(B)** DHN1 CW1, and **(C)** DHN1 W3. The calculated mass from sequence were 16955.33 Da, 17010.43, and 16978.37, respectively.

Table 3.1: Comparison of DHN1 and its mutants molecular weight determined by mass spectrometry with their calculated molecular weight.

Protein name	**Observed MW form MALDI-TOF (Da)**	**Expected MW form sequence (Da)**
DHN1	16951.45	16955.33
DHN1 CW1	16965.29	17010.43
DHN1 W3	16977.01	16978.37

TH-3004_146106002

3.6 Characterization of DHN1 and its mutant as an IDP:

3.6.1 Amino acid compositions profile and disordered prediction plots

To obtain better insights into the disorder propensity of DHN1 and its mutants DHN1 CW1 and DHN1 W3, an analysis of their amino acid composition was conducted using the composition profiler tool. **Figure 3.8** shows the relative enrichment in the amino acid composition of DHN1, DHN1 CW1, DHN W3 and an ordered protein Hen Egg White Lysozyme (HEWL) compared with SwissPort version 51 protein databases. The result obtained from the amino acid composition analysis revealed enrichment of DHN1, DHN1 CW1, and DHN1 W3 in disorder-promoting amino acid residues (Pro, Ser, Gly, and Gln)) and depleted in order-promoting residues (Leu, Val, Cys, Tyr, Trp, and Ile). While an

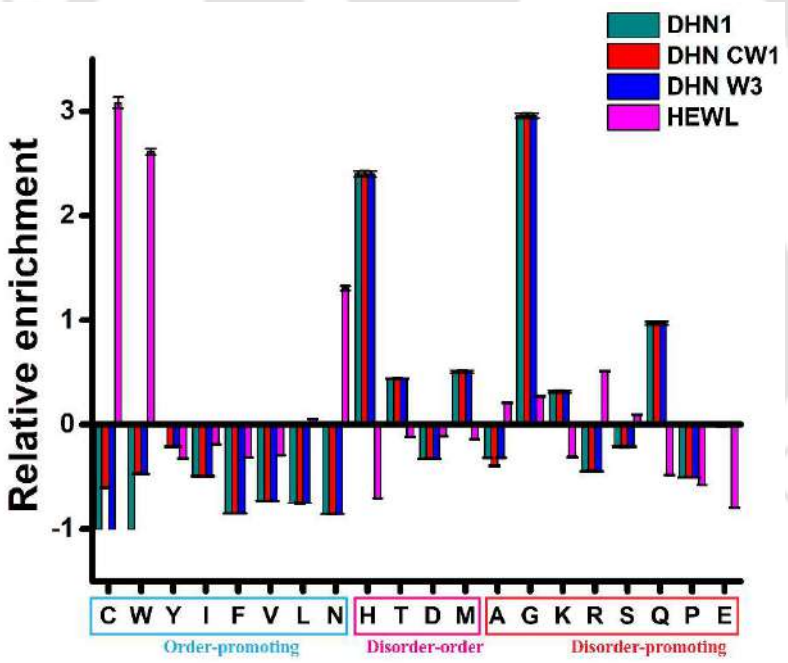


Figure 3.8 Amino acid composition analysis. Comparison of amino acid composition profile of DHN1 its mutants DHN1 CW1, DHN1 W3 with ordered protein, Hen Egg White Lysozyme (HEWL).

ordered protein, HEWL is relatively enriched in order promoting residues, including bulky hydrophobic (Ile, Leu, and Val) and aromatic residues (Phe, Tyr, and Trp), which are mainly involved in the hydrophobic core formation of well-folded globular proteins. Amino acid

TH-3004_146106002

composition profile of DHN1, DHN1 CW1, and DHN1 W3 points towards their intrinsic disordered state in native form as, they are enriched in charged and deficient in hydrophobic amino acid residues[22,25,33].

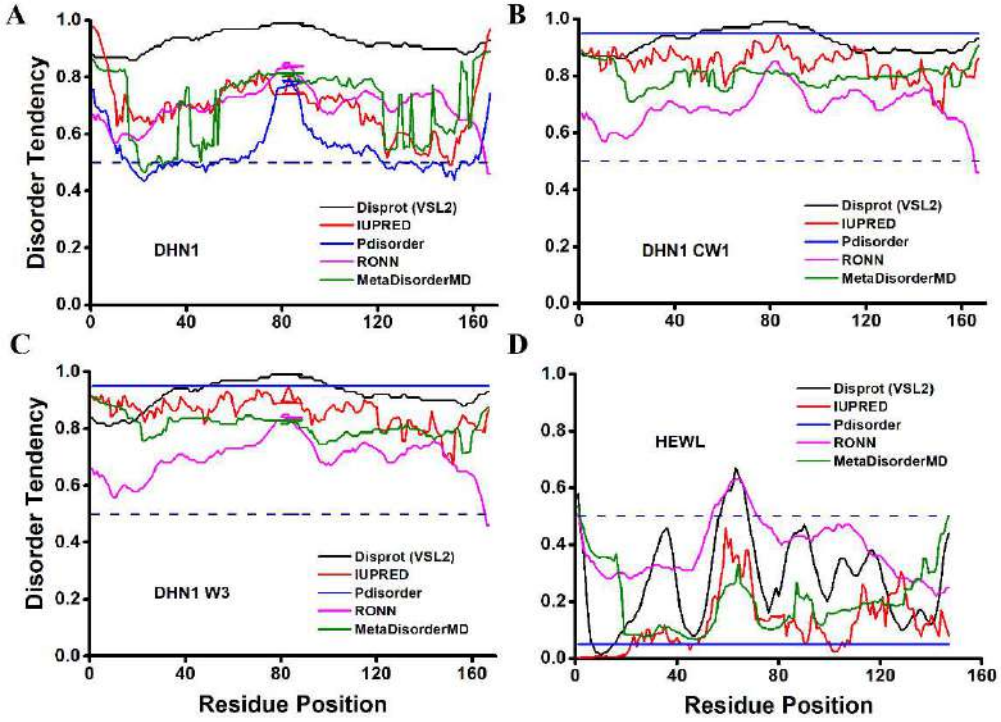


Figure 3.9 Disorder plot of DHN1 and its mutant. Comparison of intrinsic disorder prediction plots of **(A)** DHN1; **(B)** DHN1 CW1; **(C)** DHN1 W3; and **(D)** HEWL obtained by various predictor.

Further, intrinsic disorder prediction profiles of DHN1, DHN CW1, and DHN1 W3 were analyzed using the online Genesilico MetaDisorder server. **Figure 3.9A, B, C, and D** displayed the disorder prediction plot for DHN1, DHN1 CW1, DHN1 W3, and HEWL, respectively using the different predictors like Disprot (VSL2), IUPRED, Pdisorder, RONN, and MetaDiorderMD. Every amino acid of DHN1, DHN1 CW1, and DHN1 W3 shows the disorder tendency value above 0.5. DHN1, DHN1 CW1, and DHN1 W3 shows the similar but not identical disorder profile. A thorough inspection of amino acid sequences using various disorder predictor tools evidently predicted that the full length of DHN1, DHN1 CW1, and DHN1 W3 is fully disordered.

TH-3004_146106002

3.6.2 Charge Hydropathy plot and hydrophobic cluster analysis

To further explore the structural disorderness in the DHN1 and its mutants DHN1 CW1 and DHN1 W3, we calculated the mean net charge <R> and the mean hydrophobicity <H> using the PONDR server. A plot of mean net charge versus mean hydrophobicity (the Charge Hydropathy plot) gives a clear separation between the structured and disordered proteins into distinct regions. The Charge Hydropathy plot (CH-plot) based on the approach that the folding of a protein is governed by a balance between attractive forces (e.g., hydrophobic interactions) and repulsive forces (Coulomb or electrostatic repulsion)[14]. **Figure 3.10** displays the charge hydropathy plot of DHN1, DHN1 CW1, and DHN1 W3.

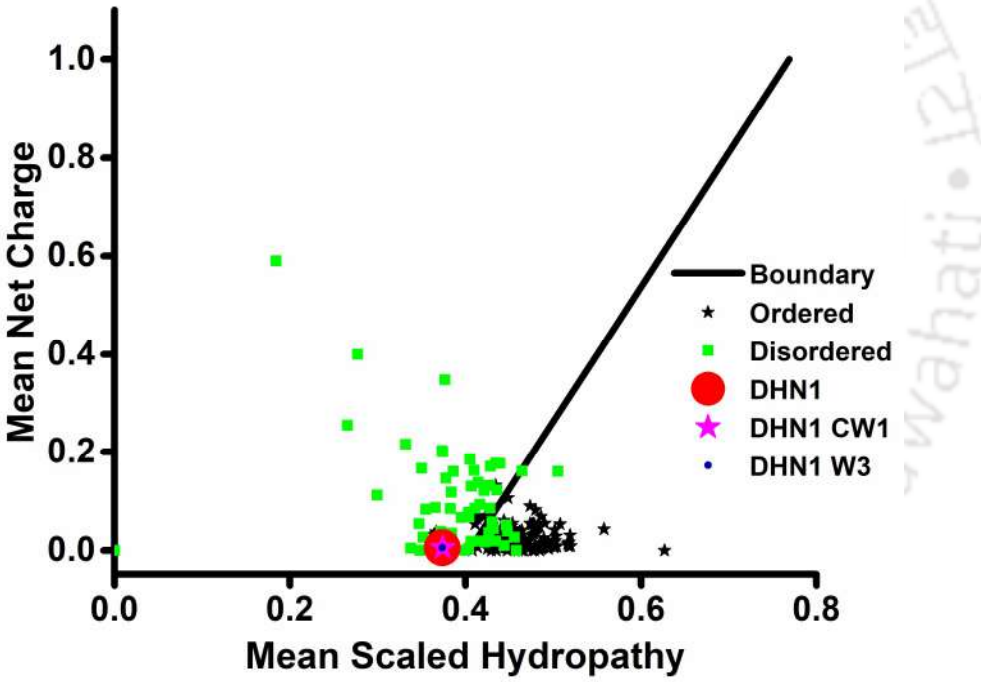


Figure 3.10. Charge hydropathy plot for DHN1 and its mutants. Charge-hydropathy plot for DHN1 (red filled circle); DHN1 CW1 (Magenta filled star); and DHN1 W3 (Blue filled circle).

TH-3004_146106002

The CH-plot is divided into two separate regions by a boundary. On the right-hand side of boundary, a set of structured proteins (stars) are present, which have low net charge and high hydrophobicity. On the left-hand side of boundary, a set of intrinsically disordered proteins (filled squares) are present, which have high net charge and low mean hydrophobicity. This result indicates that the extended IDP (native coils or native pre-molten globules) occupy unique region of the charge-hydropathy phase space and that they are reliably separated from the compact globular proteins by a simple linear boundary. DHN1, DHN1 CW1, DHN1 W3 clustered on left side of boundary with the unstructured proteins. Interestingly, the single and double amnio acid mutation does not affect the mean net charge and mean hydrophobicity of DHN1 protein. This result suggests the intrinsically disordered nature of DHN1 and its mutants DHN1 CW1 and DHN1 W3.

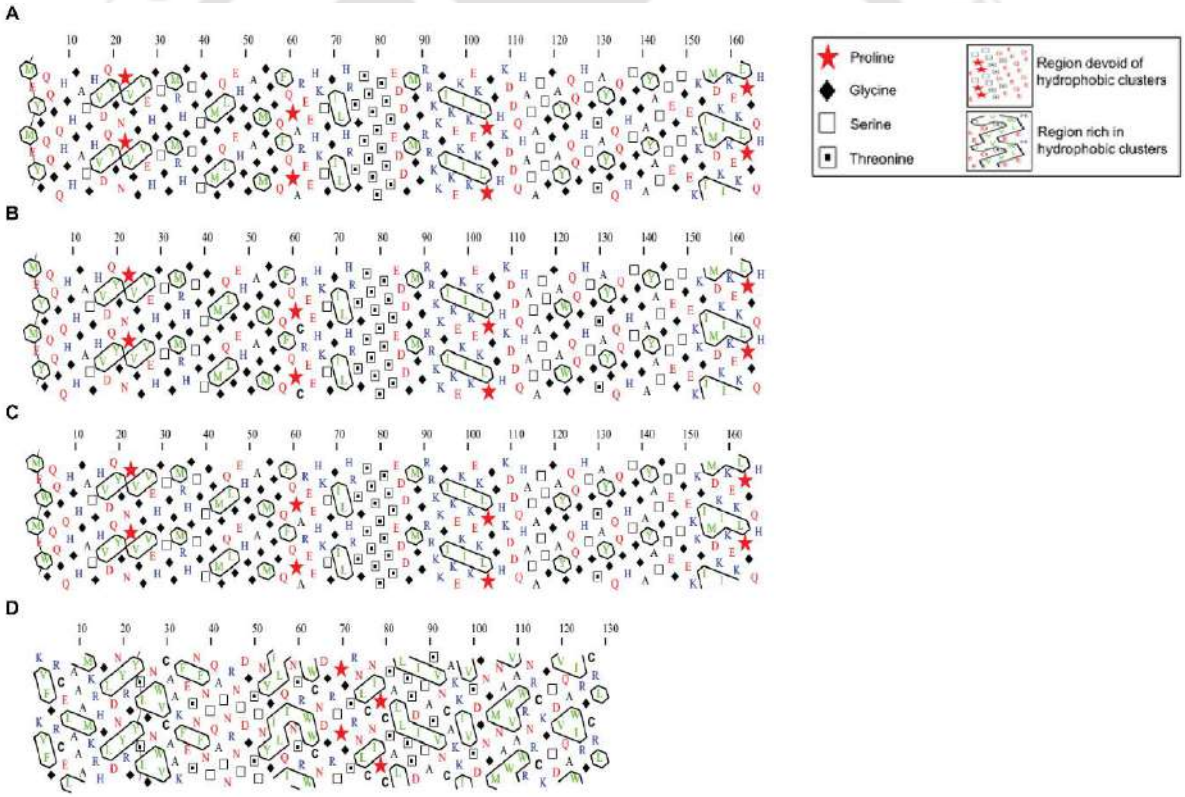


Figure 3.11 Comparison the hydrophobic cluster analysis of **(A)** DHN1; **(B)** DHN1 CW1; **(C)** DHN1 W3 with order protein **(D)** HEWL determined by using DRAWHCA1.0.2 program**.**

TH-3004_146106002

Further, the hydrophobic cluster analysis (HCA) was done using DRAWHCA 1.0.2 program for visual illustration of hydrophobic pates in DHN1, DHN1 CW1, and DHN1. Hydrophobic cluster analysis gives the information of folding repositories of the protein sequence. **Figure 3.11A, B, C and D** depicts the HCA plot of DHN1, DHN1 CW1, DHN1 W3, and HEWL, respectively. The HCA plot delineate the hydrophobic regions in DHN1 protein. Detailed investigation reveals that the DHN1, DHN1 CW1, and DHN1 W3 are depleted in the hydrophobic regions. DHN1 and its mutants DHN1 CW1 and DHN1 W3 indicating the similar HCA profile, indicating that the substitution of amino acids in DHN1 protein does not change the hydrophobic regions. Conclusively, the depletion of the hydrophobic clusters also hints towards the structural disorder in DHN1, DHN1 CW1 and DHN1 W3.

3.6.3 Anomalous SDS-PAGE mobility and heat resistance

To experimentally confirm the extent of disorder; DHN1, DHN1 CW1 and DHN1 W3 were run on reducing SDS-PAGE. **Figure 3.6A, B and C** shows the anomalous behavior of DHN1 and its mutant on SDS-PAGE. DHN1 and its mutants DHN1 CW1 and DHN1 W3 show an apparent MW of ~19 kDa on reducing SDS-PAGE, however, their actual mass and expected mass is 16.9-17 kDa. Such anomalous mobility of IDPs/IDRs is not uncommon. This retarded mobility appears due to abnormally low SDS- binding to hydrophilic residue in the IDPs/IDRs and also due to excess charge in protein. Such anomalous behavior of DHN1, DHN1 CW1 and DHN1 W3 also indicates towards the intrinsic disorder in native state.

Further, upon heating DHN1, DHN CW1 and DHN1 W3 up to 95 °C for 1 hour. No changes observed in the band intensity and MW DHN1, DHN1 CW1, and DHN1 W3 hinting towards their heat resistance property. This heat-resistance property of DHN1 is one of the characteristic features of IDPs/IDRs.

3.6.4 Absorption spectra of DHN1 and its mutants

The absorption spectra of purified DHN1, DHN1 CW1 and DHN1 W3 were recorded to ascertain their identity and purity, spectroscopically. **Figure 3.12** illustrate the absorption spectra of DHN1, DHN1 CW1 and DHN W3. The DHN1 protein contains a single

TH-3004_146106002

phenylalanine residue, and five tyrosine residues show the distinct absorption profile of tyrosine. However, the existence of tryptophan in DHN1 CW1 and DHN1 W3 is confirmed as it displayed a distinctive strong absorption profile of tryptophan and caused the peak broadening. The shoulder peak at 280 nm due to tryptophan absorption is clearly visible in DHN1 CW1 and DHN1 W3. Further, the dominance of four tyrosine residues is evident in the absorption spectra of DHN1 CW1 and DHN W3. Inset of **Figure 3.12** show the existence of absorption in DHN1 and its mutants beyond the 320 nm extending till upto 800 nm. These absorption beyond 320 nm arising from the ProCharTS, discussed in upcoming chapter.

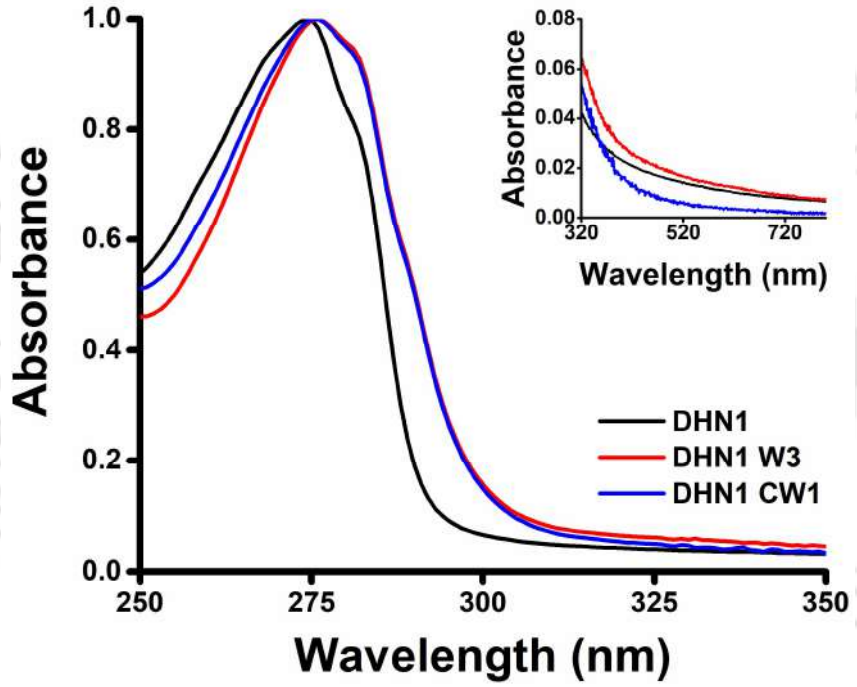


Figure 3.12. Absorption spectra of DHN1, DHN1 CW1 and DHN1 W3 dissolved in deionized water. In inset spectrum shown for longer wavelength (320-800 nm) of DHN1, DHN1 CW1 and DHN W3 in expanded form, respectively.

TH-3004_146106002

3.6.5 Steady state fluorescence and anisotropy of DHN1, DHN1 CW1 and DHN1 W3

Intrinsic UV-visible fluorophores existing in the side chain of aromatic amino acids like tryptophan and tyrosine have been extensively used to study the structure and dynamics of several proteins owing to their spectral sensitivity toward the local environment[83,191–193]. Amongst the three intrinsic fluorophores of proteins, tryptophan residue is the most frequently used to uncover structural dynamics. Several studies have been reported in context to understand the correlation drawn between the position of the maximum emission

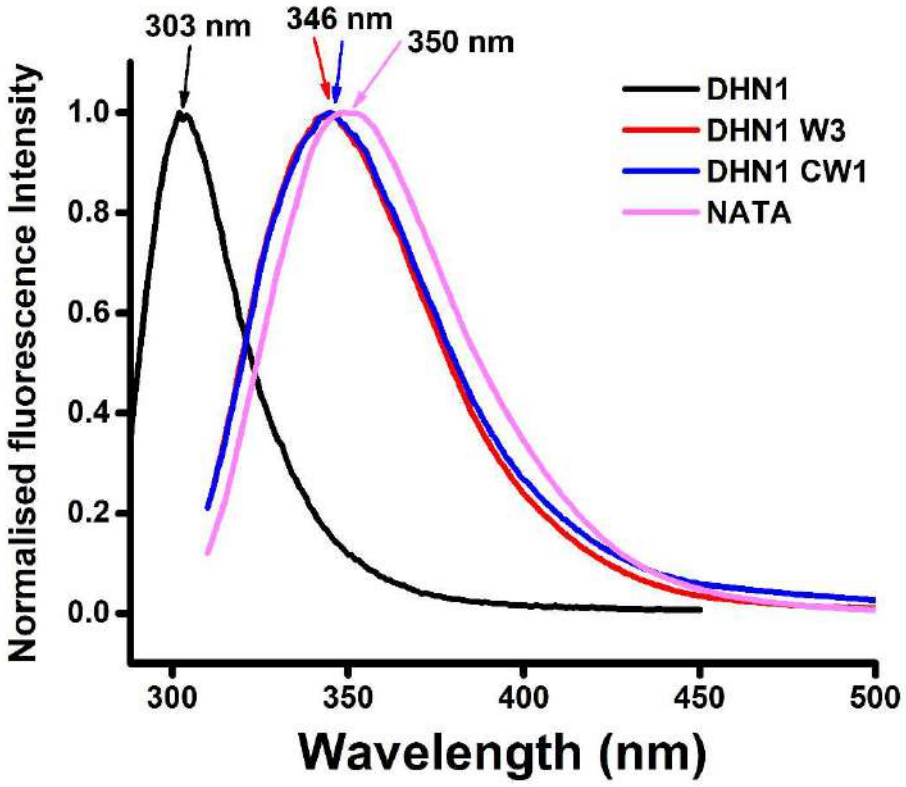


Figure 3.13. Steady state fluorescence emission. Normalized steady state fluorescence of 20 µM DHN1 tyrosine (black), DHN1 CW1 tryptophan (blue), DHN1 W3 tryptophan(red), and 15 µM NATA (magenta) dissolved in the 50 mM Sodium Phosphate buffer (pH 7.4), containing 5 mM TCEP in case of DHN1 CW1. Excitation wavelength 274 nm and 295 nm used for excitation of chromophore present in side chain of tyrosine and tryptophan or NATA, respectively. The slit width kept for excitation and emission are 2 nm and 5 nm, respectively.

TH-3004_146106002

wavelength (λ_{max}) of tryptophan in a protein and its exposure to the solvent. Fluorescence emission from tryptophan is highly sensitive to the local environment and shows a blue shift when it is buried inside the protein's hydrophobic core and red shift when exposed to the aqueous solvent[156]. Fluorescence emission of DHN1 tyrosine and DHN1 CW1, DHN1 W3 tryptophan were performed by exciting sample at 274 nm and 295 nm, respectively. **Figure 3.13** illustrates the steady state fluorescence spectra of DHN1 tyrosine, and DHN1 CW1 and DHN1 W3 tryptophan along with a tryptophan derivative known as NATA (N-acetyl tryptophan amide). A fluorescence emission peak observed for DHN1 at 303 nm, DHN1 CW1 and DHN1 W3 tryptophan show emission peak at 346 nm, while NATA shows emission maxima at 350 nm. It is well-established that the fully solvent-exposed tryptophan in protein shows its emission maxima at 350 nm[179,194,195].

Table 3.2 Steady state fluorescence anisotropy of NATA; and DHN1 CW1, DHN1 W3 tryptophan.

Sample	**Steady state anisotropy (r_{ss})**
NATA	0.004 [0.001][a]
DHN1 CW1	0.050 [0.005]
DHN1 W3	0.054 [0.002]

[a] The number in square brackets indicates the standard deviation for n= 2−3.

The emission maxima of DHN1 CW1 and DHN1 W3 tryptophan at 346 nm suggests the predominant exposure of tryptophan to the aqueous solvent. Further, to obtain insights into the structure and rotational dynamics of DHN1 CW1 and DHN1 W3, steady state fluorescence anisotropy (r_{ss}) of its tryptophan were analyzed at room temperature. The steady state anisotropy (r_{ss}) is served as an excellent technique to investigate the rotational dynamics of the fluorophore. The steady state anisotropy (r_{ss}) provides information about the rotational freedom of tryptophan residue[196]. The recorded anisotropy values of the DHN1 CW1 and DHN1 W3 tryptophan **(shown in Table 3.2)**, suggest fast rotation of DHN1 CW1 and DHN1 W3 tryptophan as it can rotate more freely. DHN1 CW1 and DHN1 W3 tryptophan shows the similar but not identical fluorescence properties. Conclusively, the predominant exposure to solvent and the fast rotational motion of DHN1 CW1 and DHN1 W3 tryptophan clearly suggest that both proteins have an open and loose structure

TH-3004_146106002

in the vicinity of tryptophan. This result suggests that the DHN1 protein has the open structure at N-terminal and near the conserved K-segment.

3.6.6 Time-resolved fluorescence lifetime of DHN1 CW1 and DHN1 W3

To unveil the structural information of the DHN1 CW1 and DHN1 W3 proteins, nanosecond time-resolved fluorescence lifetime (τ) of their tryptophan were monitored. The sample was excited using 295 nm DeltaDiode pulsed light source, and emission were

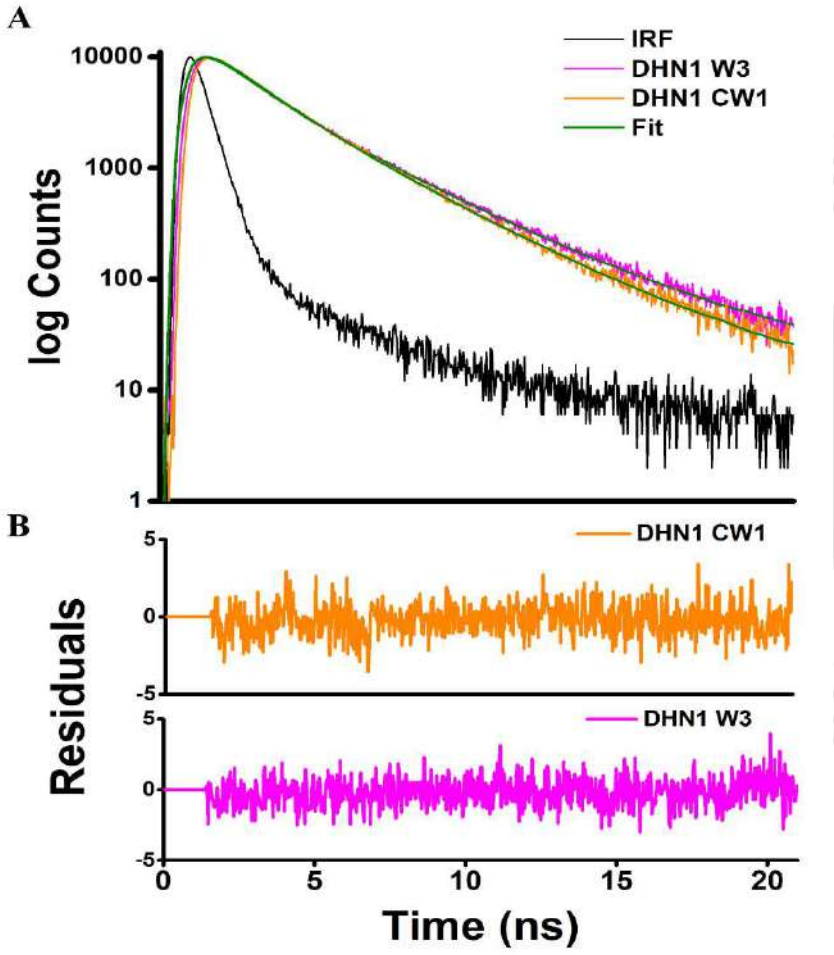


Figure 3.14 Time-resolved fluorescence intensity decay profile of 20 μM DHN1 CW1 and DHN1 W3 dissolved in the 50 mM sodium phosphate buffer (pH 7.4), containing 5 mM TCEP in case of DHN1 CW1. Panel shows **(A)** Time-resolved fluorescence intensity decay fit of DHN1 CW1 and DHN1 W3, and **(B)** Residual for fit.

TH-3004_146106002

Table 3.3 Time-resolved fluorescence lifetime values of DHN1 CW1 and DHN1 W3 tryptophan obtained from discrete fit.

Protein	α_1	τ_1 (ns)	α_2	τ_2 (ns)	τ_m (ns)	χ^2_R
DHN1 CW1	0.44	2.96 [0.014][a]	0.55	1.51 [0.016]	2.15 [0.004]	1.02
DHN1 W3	0.42	3.22 [0.013]	0.58	1.39 [0.002]	2.14 [0.003]	1.05

[a] The number in the square bracket indicates the standard deviation for n=2−3

collected with longpass filter WG320 at magic angle 54.7°. The fluorescence lifetime is the measure of the average amount of time a fluorophore stays in its excited state prior to coming back to the ground state. **Figure 3.13A** illustrate the time-resolved fluorescence intensity decay profile of DHN1 CW1 and DHN1 W3 tryptophan in pH 7.4 buffer at room temperature. The lifetime values of DHN1 CW1 and DHN1 W3 tryptophan obtained after fitting the intensity decay profile are shown in **Table 3.3**

The analysis of time-resolved fluorescence decay revealed the existence of two lifetimes in DHN1 CW1 and DHN1 W3. In DHN1 CW1 large component of 2.95 ns with amplitude of 0.44 and a short component of 1.51 ns with amplitude of 0.55 was present. Similarly, in DHN1 W3 large component of 3.21 ns with amplitude of 0.41 and a short component of 1.39 ns with amplitude of 0.58 was present. The observed mean fluorescence lifetime of DHN1 CW1 and DHN1 W3 tryptophan were 2.15 and 2.14 ns, respectively. The possible reason for two-exponential decay of DHN1 CW1 and DHN1 W3 tryptophan may be due to the presence of tryptophan rotamers or the existence of DHN1 CW1 and DHN1 W3 in two conformation or dynamic processes occurring during the excited-state lifetime. Further, the richness of charged amino acid residues in DHN1 CW1 and DHN1 W3 could influence the fluorescence decay of tryptophan (Detailed explanation given in forthcoming chapter 5).

TH-3004_146106002

3.6.7 Secondary structure determination of DHN1, DHN1 CW1 and DHN1 W3 by CD

To explore the secondary structure content of DHN1 and its mutants DHN1 CW1 and DHN1 W3, far-UV Circular Dichroism (CD) spectra of DHN1, DHN CW1 and DHN1 W3 in the range of 190 to 260 nm were recorded. **Figure 3.15** shows the far-UV CD spectra of DHN1, DHN1 CW1 and DHN1W3 at room temperature. The DHN1, DHN CW1 and DHN1 W3 show a negative ellipticity peak at 198 nm; which completely resembles with the spectrum of unstructured proteins. The similar secondary structure content of DHN, DHN1 CW1 and DHN1 W3 suggests there is no structural alteration due to the incorporation of extra tryptophan and cysteine residue in DHN1. The CD spectra of DHN1, DHN1 CW1 and DHN1 W3 were analyzed using DichroWeb server and CDSSTR program was used to compute their secondary structure content.

Table 3.4 depicts the secondary structure content in DHN1, DHN1 CW1 and DHN1 W3. DHN1, DHN1 CW1 and DHN1 W3 show the highest percentage of unordered content in their secondary structure and clearly point towards the random coil conformation of DHN1 protein. The experimental curve was well superimposed on reconstructed curve (fitted CD curves are shown in **Figure A1** of the Appendix).

TH-3004_146106002

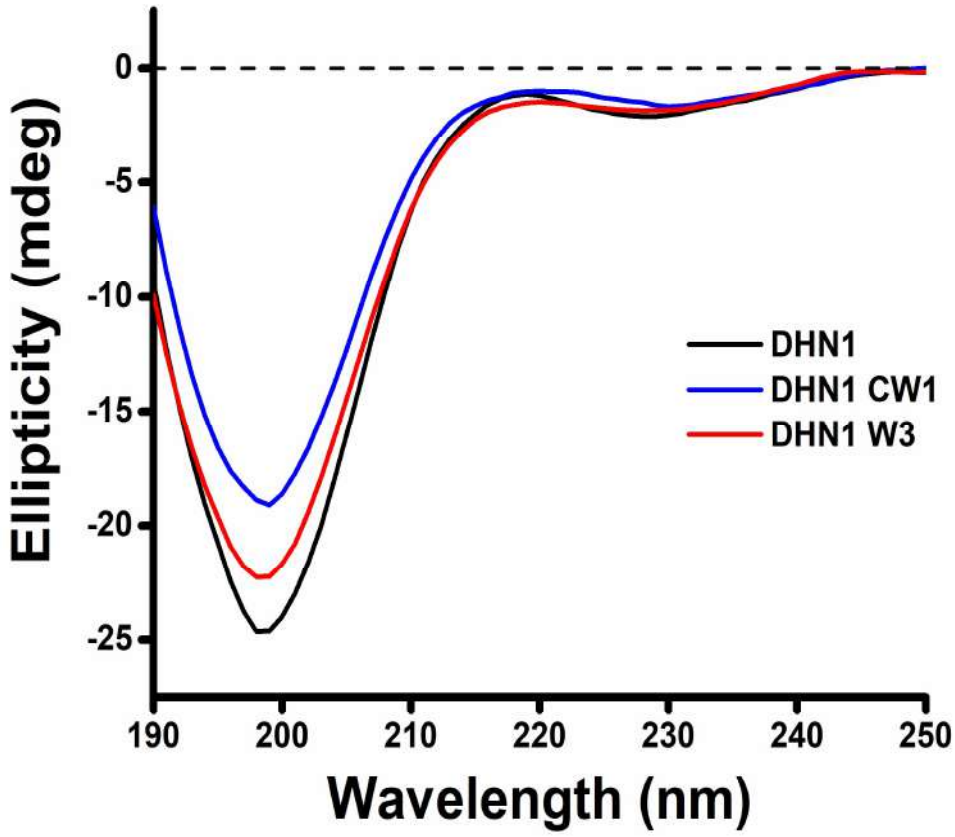


Figure 3.15 Far-UV CD spectra of 7 μM DHN1, DHN1 CW1, and DHN1 W3, dissolved in water, 1mM TCEP in case of DHN1 CW1.

Table 3.4 Secondary structure content of DHN1, DHN1 CW1 and DHN1 W3 analyzed by DichroWeb

Protein	Helix %	Strands %	Turn %	Unordered %
DHN1	3 [1.155]	24 [4.243]	16 [3.916]	56 [9.0921]
DHN1 CW1	3.5 [0.577]	23 [2.828]	15 [2.449]	57.5 [5.568]
DHN1 W3	2.5 [0.701]	22.5 [3.535]	14 [2.828]	59 [7.072]

[a] The values in square bracket shows the standard deviation for n=2−3.

TH-3004_146106002

3.7 Conclusions

- The incorporation of Trp-Cys pair and single Trp in DHN1 at desired site was done successfully using site-directed mutagenesis; generating two mutants namely DHN1 CW1 and DHN1 W3.

- DHN1 and its mutant DHN1 CW1, DHN1 W3 were expressed successfully in BL-21 Codon^{+} (DE3) strain cells by inducing with 0.25 mM IPTG.

- DHN1 and its mutant DHN1 CW1, DHN1 W3 were purified (2−3 mg/L of the media) and molecular mass detected successfully.

- The amino acid composition profile, charge hydropathy plots and the hydrophobic cluster analysis illustrate that DHN1 and its mutants DHN1 CW1 and DHN1 W3 are fully disordered in native state and have low hydrophobicity.

- Anomalous mobility on reducing SDS-PAGE and the heat resistance points towards the disorderedness in DHN1 protein.

- The fluorescence emission of DHN1 CW1 and DHN1 W3 shows the free rotation and solvent exposure of its tryptophan indicating that DHN1 CW1 and DHN1 W3 are unstructured.

- DHN1 and DHN1 CW1 an DHN1 W3 share the similar secondary structure contents showing the minimal effect of incorporation of tryptophan and cysteine residue in DHN1 protein.

TH-3004_146106002

Chapter 4

Probing the structural transitions in DHN1 protein by new tool: ProCharTS

TH-3004_146106002

4.1 Introduction:

Investigation of biomolecules without any perturbation is the sole requirement to understand their structure and dynamics in natural state. Intrinsic UV-visible chromophores existing in the side chain of aromatic amino acids like tryptophan and tyrosine have been extensively used for such purpose in several proteins[125,156,167,178–180]. The absorbance and fluorescence signature of aromatic amino acids in proteins are also well characterized [183,197].

However, several important protein classes like intrinsically disordered proteins (IDPs) are rich in charged and devoid of aromatic amino acid residues in their sequence[16,33,198]. These impose several constraints on monitoring the structural changes in IDPs[198]. To overcome these limitations, there is thus a demanding need for another intrinsic chromophores to study the structure and dynamics of IDPs. Unstructured proteins and unstructured regions in the proteins can be characterized by a various biophysical and biochemical methods such as small angle X-ray scattering, Raman optical activity, circular dichroism (CD) and protease sensitivity to name a few. Here, we used the newly discovered tool Protein Charge Transfer Spectra (ProCharTS) to monitor the structural transitions in an IDP protein.

ProCharTS originates due to photoinduced charge transfer from: polypeptide backbone to NH_3^+ in the lysine; COO^- in glutamate to polypeptide backbone; and COO^- in glutamate to NH_3^+ in lysine. The absorption intensities of ProCharTS in 250-800 nm range are dependent on the 3D spatial proximity of lysine/glutamate residue[144]. IDPs are rich in charged amino acid residues and can thus be studied by ProCharTS in native state.

Dehydrin (DHN1) protein from *Zea mays* contains about 29% charged amino acids with the abundance of lysine, glutamate and histidine. It will be an appropriate approach to employ the ProCharTS as a new label-free spectroscopic tool to monitor the structural transitions in DHN1. In this chapter, we take advantage of charge richness in DHN1 to sense their structural transitions using ProCharTS. We investigated the conformational changes in DHN1 induced by the change in pH and temperature of aqueous medium by ProCharTS and CD spectra.

Our result revealed that in presence of tyrosine, ProCharTS absorbance was substantially decreased, specifically at wavelength where absorption by tyrosine was near its maximum.

TH-3004_146106002

Significant changes in the ProCharTS spectrum were observed with changing pH in the range of 3—11. ProCharTS intensity was sensitive to the temperature-induced structural transitions in DHN1 between 25-85°C. Further, the presence of high salt (250 mM KCl and NaCl) in the medium also affects the ProCharTS spectrum. In conclusion, we proposed the utility of ProCharTS as a new label-free intrinsic probe to monitor the structural transitions in IDPs.

TH-3004_146106002

4.2 Result and Discussion:

4.2.1 Protein Charge Transfer (ProCharTS) Absorption Spectra of DHN1

The amino acid sequence of DHN1 protein (shown in Figure 3.2 of Chapter 3) shows the abundance of charged amino acid residues. DHN1 is rich in lysine, glutamate and aspartate with few sparsely distributed tyrosine and phenylalanine residues.

To investigate the ProCharTS absorption of DHN1 protein, absorption spectrum was measured in UV-visible range of 250 – 800 nm. **Figure 4.1** shows the absorption spectrum obtained for the DHN1 in deionized water. The molar extinction coefficient of DHN1 ProCharTS is shown in **Table 4.1**. From the inset in **Figure 4.1**, it is noticeable that the effect of tyrosine absorption in DHN1 is diminished after 300 nm. It is clear that stronger absorption from 5 Tyr in the DHN1 can significantly mask the weaker ProCharTS spectrum in same protein between 250 and 300 nm. Further, the weak chromophore in phenylalanine has negligible influence on the DHN1 ProCharTS spectrum at 257 nm. Hence the absorption intensity arising from ProCharTS can be explicitly identified only beyond these wavelengths as shown in **Figure 4.1** when all other amino acid chromophores are silent. Thus, the molar extinction coefficient observed at and beyond 300 nm arises exclusively from ProCharTS. DHN1 reveal a distinctive broad tail of ProCharTS absorption between 300—800 nm as observed earlier with α_3C and PEST fragment[144,199]. Inset of **Figure 4.1** reveal that when the strong contribution from tyrosine is deducted from DHN1 spectrum near 250—290 nm using the molar extinction coefficient reported previously[200], a weakened spectrum of ProCharTS is left behind. The additional absorption contributed by ProCharTS in the region of 250-300 nm, where the aromatic amino acids residue absorbs, if unaccounted, can lead to substantial overestimation of protein concentration based on aromatic amino acid absorption. However, at wavelengths beyond 300 nm (315, if tryptophan present in protein), the ProCharTS spectra are devoid of any contribution from aromatic amino acid.

To evaluate the contribution of intramolecular or intermolecular interaction in absorption spectra of DHN1, absorbance of DHN1 dissolved in deionized water with their increasing concentration (25—150 μM) was collected. **Figure 4.2** demonstrate the absorbance of

TH-3004_146106002

DHN1, which shows the linear increase in DHN1 absorption with increasing concentration at multiple wavelengths. This absorption linearity of DHN1 clearly eliminates the possibilities of ProCharTS spectra arising due to any intermolecular interactions.

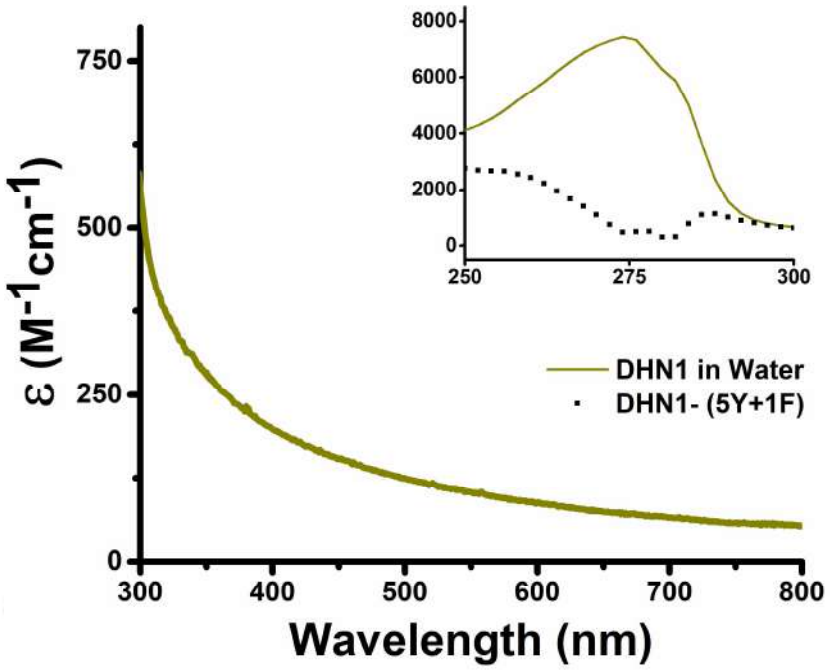


Figure 4.1 Absorption Spectra of DHN1 protein in deionized water. In inset spectrum shown for shorter wavelength (250-300 nm) of DHN1 in expanded form. Spectrum left behind after subtraction of aromatic chromophore is also shown in dotted line in inset. Extinction coefficient reported are an average of multiple concentration between 50—100 μM.

Table 4.1: ProCharTS absorption: Molar extinction coefficient of absorption arising from ProCharTS, in DHN1 is displayed. Other conditions are similar to **Figure 4.1.** The numbers in brackets indicate standard deviations for n = 2–3.

DHN1	Charged residue fraction	ε (M^{-1} cm^{-1}) at different wavelengths in nm												
		315	**320**	**325**	**350**	**400**	**450**	**500**	**550**	**600**	**650**	**700**	**750**	**800**
	29%	396 (72)	370 (68)	351 (64)	284 (54)	198 (34)	154 (21)	125 (15)	103 (10)	87 (8)	74 (5)	65 (4)	57 (4)	53 (5)

TH-3004_146106002

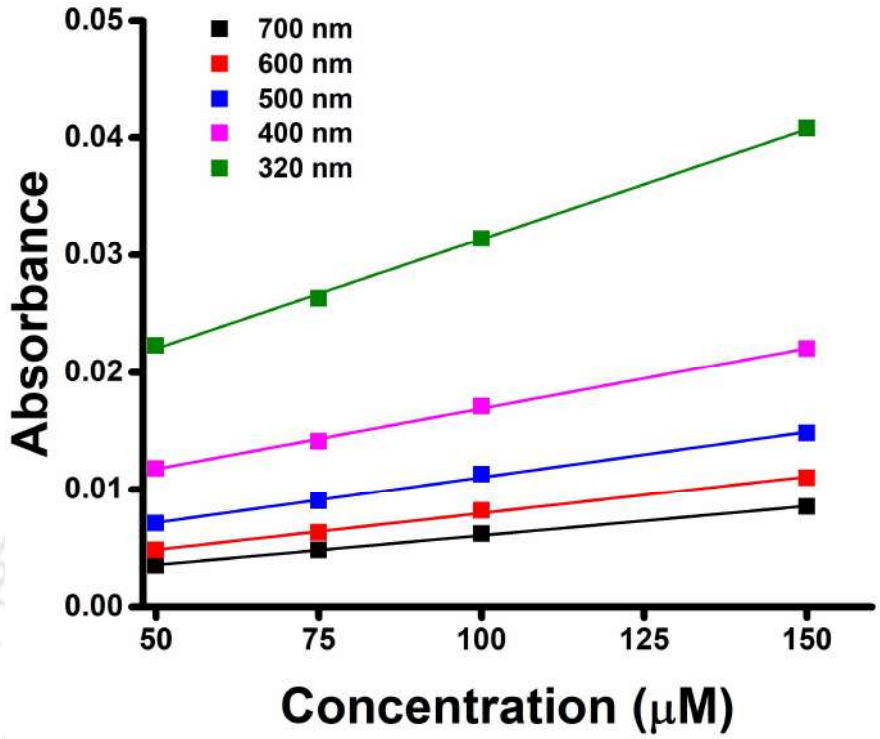


Figure 4.2 Variation of ProCharTS absorbance with protein concentration for DHN1 at chosen wavelengths.

4.2.2 ProCharTS compared with simulated Rayleigh Scatter

To better understand whether the observed spectral feature at and after 300 nm arise due to Rayleigh scattering, DHN1 absorption spectrum was compared with simulated Rayleigh scatter. The simulated Rayleigh scatter was acquired by plotting the obtained absorption spectrum of DHN1 as λ^4 dependence using **equation 4.1**[201]. In the presence of any scattering component, DHN1 absorption spectrum should overlap with the simulated Rayleigh scatter profile in the sample, which follows ($1/\lambda^4$) dependence [202,203].

TH-3004_146106002

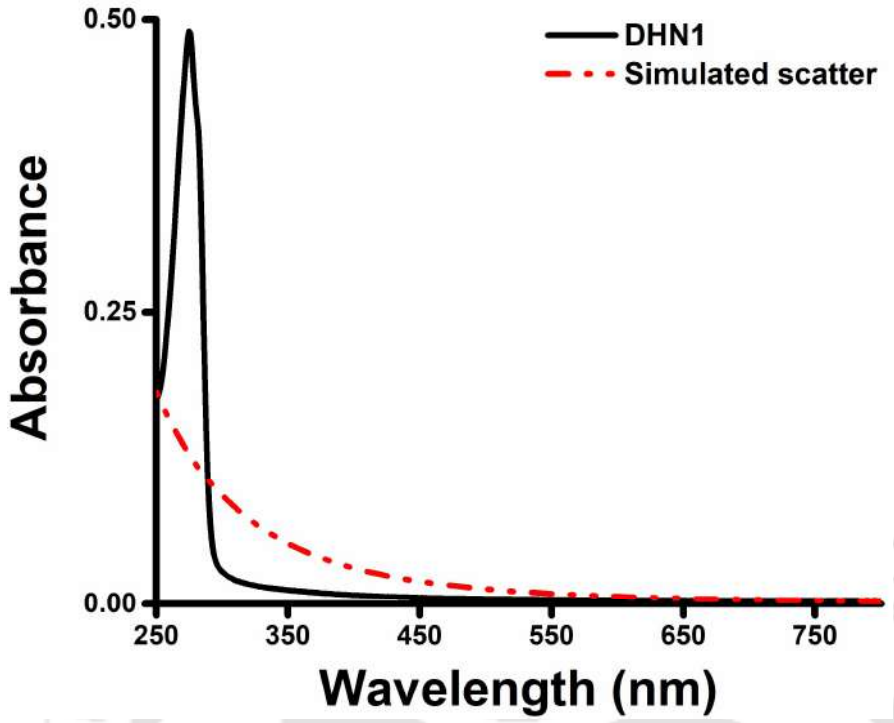


Figure 4.3 Comparison of absorbance spectra with simulated Rayleigh scatter (using $1/\lambda^4$ dependence) for DHN1.

$$\sigma_s = \frac{2\pi^5}{3}\frac{d^6}{\lambda^4}\left(\frac{\eta^2-1}{\eta^2+2}\right)^2 \qquad 4.1$$

Here, σ_s represent the Rayleigh scattering cross-section, λ is the wavelength of light, d is the diameter of the particle and $\eta = \frac{\eta_{particle}}{\eta_{medium}}$ is the relative refractive index.

The comparison of absorption spectra of DHN1 with simulated Rayleigh scatter profile is shown in **Figure 4.3**. Poor overlay of ProCharTS spectra with simulated Rayleigh scattering profile pointing towards the absence of any scattering artifact in the observed absorption spectra of DHN1 from 250—800 nm. It is interesting to study the utility and sensitivity of ProCharTS to monitor the structural transitions in an IDP, DHN1.

Several exercises were performed to assess the sensitivity of ProCharTS spectrum with structural transitions in the DHN1. In the first, the pH of the medium was varied from pH

TH-3004_146106002

3 to 11, so as to both alter the ionization state of charged amino acid and moderately perturb the structure of the protein, as discussed in the following section.

4.2.3 Effect of pH

To illustrate the effect of pH on the ProCharTS of DHN1 protein, the absorption spectra of DHN1 was recorded from 250 to 800 nm at varying pH (3–11). **Figure 4.4A** shows the effect of changing pH on ProCharTS absorption spectra of DHN1. The inset of **Figure 4.4A** shows the change in absorption arising from predominantly five tyrosine residues in DHN1 with a peak near 275 nm. **Figure 4.4B** shows the percentage change in absorption intensity of DHN1 at different pH (3–11).

However, significant contribution from ProCharTS is evident here based on the spectral perturbations observed with changing pH. A prominent dip in pH 3 is observed in the spectrum between 250 – 290 nm, while further decline is visible in pH 11 in comparison to spectra in pH 5 – 9. There is no change observed at 292 nm in pH 11, where tyrosinate anion if present can increase the absorption significantly. In the main **Figure 4.4A**, DHN1 shows moderately higher ProCharTS absorption in deionized water as compared with different pH unlike other protein (data not shown)[199]. Aside from spectrum for pH 3, spectra for remaining pH remain low in intensity and similar to each other at longer wavelengths. **Figure 4.4C** reveals minor change in pH 3 and 5 and no significant change in CD spectra at long wavelength region.

The pH effect observed with ProCharTS (**Figure 4.4 A, B and C**) can be analyzed to detect common feature. At pH 3, the carboxylate group in both glutamate and aspartate are likely to be protonated and uncharged, so all proteins at this pH are likely to have a net charge from histidine, lysine and arginine residue only. This also means that salt bridge interaction between oppositely charged residue cannot exist in pH 3. The higher ProCharTS signal appears to correlate with increased helix and strand populations along with more ordered structure compared to other pH. Interestingly, at pH 5, 7 and 9 mild changes in ProCharTS intensity were observed due to the rich presence of negative and positive charges in all charged residues in this pH range. Surprisingly, mild change in ProCharTS intensity were observed between pH 5—9 despite of having large number of histidine residues which has

TH-3004_146106002

the protonated imidazole side chain (pK_a ~6.0). In summary, pH-related changes in ProCharTS seem complex due to simultaneous changes in: a) ionization state of charged residue: b) protein conformation, and c) salt concentration. Interpretation of changes at extremes (pH 3 and 11) appear easier.

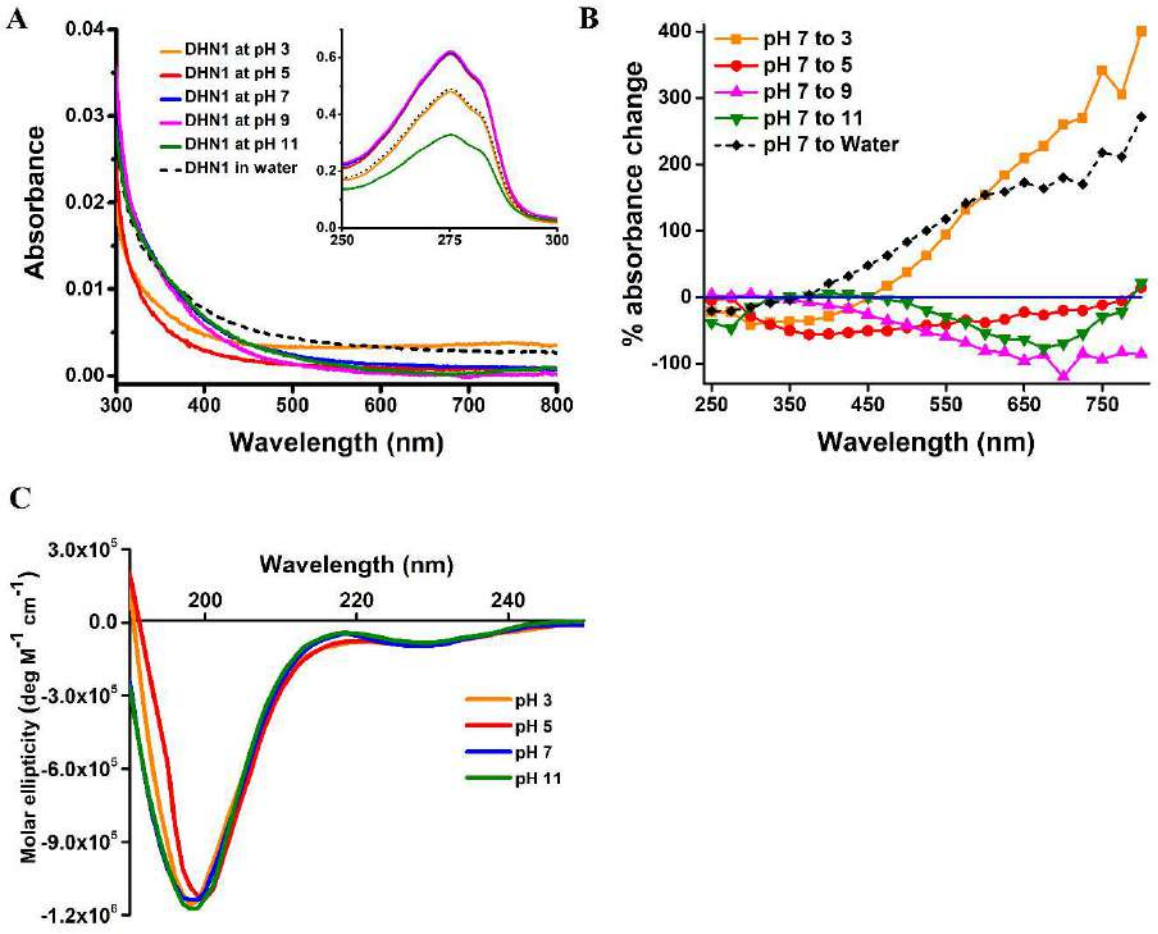


Figure 4.4 Effect of pH on ProCharTS and CD: The effect of changing pH on 125 µM DHN1 on **(A)** absorption spectra; Inset shows the absorption spectrum at shorter wavelengths **(C)** Circular Dichroism; and **(B)** % change in absorbance intensity. Percentage change in absorption intensity measured in every 25 nm at specified pH with respect to pH 7. The change in absorbance at selected wavelength were calculated as [(Absorbance at chosen pH – Absorbance at pH 7)/ Absorbance at pH7] × 100.

Thus, ProCharTS found to be sensitive to monitor the pH-induced structural transitions as shown in DHN1 protein. In the following approach, we altered the temperature of the

TH-3004_146106002

medium to disrupt the protein structure and subsequently monitored its effect on the ProCharTS spectrum.

4.2.4 Effect of Temperature

To study the effect of temperature on the ProCharTS of DHN1 protein, the absorption spectra of DHN1 was recorded from 250 to 800 nm at different temperatures (25 – 85 °C). **Figure 4.5A** shows the rise in absorption spectra of DHN1 with the increase in temperature. DHN1 absorption spectra reveal a marginal increase in absorption at higher temperature (85 and 65 °C), pointing towards minor alteration in the structure of protein.

No significant changes are seen at shorter wavelength (**Inset of Figure 4.5A**). Interestingly, spectra at 65 and 85 °C, appear nearly same, hinting at similar conformations. Upon cooling to 45 and 25 °C, the ProCharTS absorption decreases, while absorption spectrum obtained at 25 °C after cooling is significantly unlike from the spectra of the original sample (25 °C before heating). Different conformation at 25 °C pointing that upon cooling, DHN1 may have obtained another conformation from the initial one.

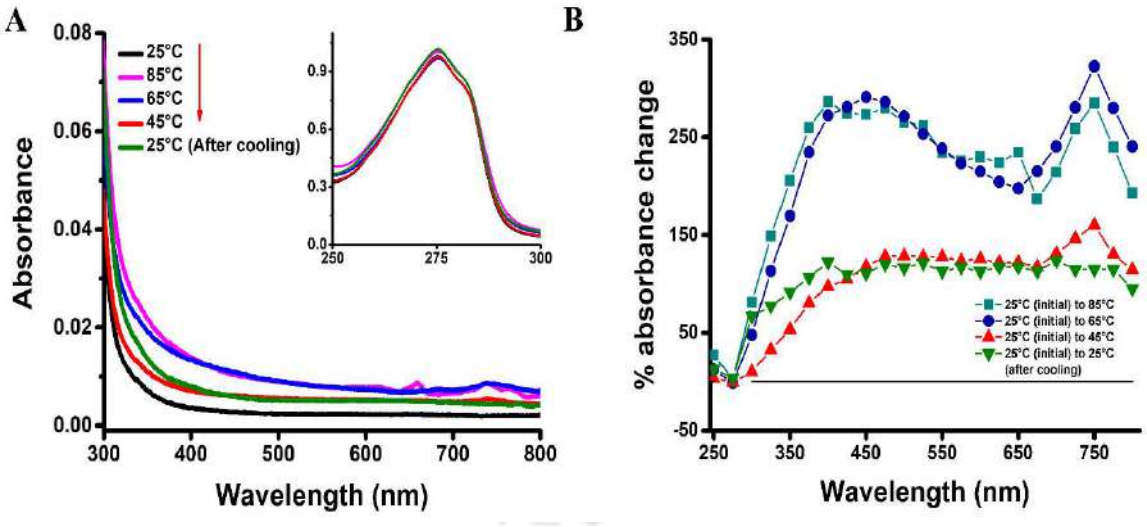


Figure 4.5 Effect of temperature on ProCharTS: The effect of temperature on **(A)** absorption spectra of DHN1 in deionized water. In inset spectrum shown for shorter wavelength (250 – 300 nm) of DHN1 in expanded form. Percentage change in absorption intensity measured every 25 nm at specified temperature for **(B)** DHN1 with respect to room temperature (25 °C). The change in absorbance at selected wavelengths were calculated as [(Absorbance at chosen temperature – Absorbance at room temperature)/ Absorbance at room temperature] × 100.

TH-3004_146106002

Overall, heat-induced changes monitored could identify structural transitions in DHN1 at 85 °C. The heat-induced structural transitions observed in DHN1 protein further supported by the circular dichroism measurement (data shown in **Figure 6.1** of chapter 6), suggests the gain of ordered structure in DHN1 at higher temperature. These acquisition of ordered structure at higher temperature was also reported previously in disordered protein[52,109].

In next section, we examine the effect of excess salt on the ProCharTS spectra of DHN1 protein. The key purpose of this experiment was to perturb the electrostatic interactions between like and opposite charges in the protein by introducing salt counter ions. As the ProCharTS spectrum is dependent on the charge localized on amino acid side chains and their interactions with those in neighboring side chain, stabilizing the charges by introducing excess salt counter ions could influence the spectrum. These changes in the ProCharTS spectra could in turn provide insight on the multiple side chain interactions contributing to the spectrum.

4.2.5 Effect of Salt

To access the effect of excess salt on the ProCharTS of DHN1 protein, the absorption spectra of DHN1 were recorded from 250 to 800 nm in presence of 250 mM NaCl and KCl. **Figure 4.6A** shows the absorption of DHN1 protein in presence of 250 mM NaCl and KCl. **Figure 4.6B** displayed the percentage change in absorption intensity of DHN1 in presence of high salt. The absorption of DHN1 is significantly decreased in the presence of NaCl and KCl at wavelengths beyond 300 nm. However, at shorter wavelength (250−300 nm) the absorption of DHN1 is substantially increased. Interestingly, the larger sodium ion appears to cause more change in ProCharTS compared to potassium in DHN1. Here, the rise and dip of ProCharTS at shorter and longer wavelengths, respectively, mirrors a similar trend

TH-3004_146106002

of ProCharTS in the presence of buffer salts compared to water (**Figure 4.4A**). Taken together, the result suggests that the salt ion can alter ProCharTS spectra significantly.

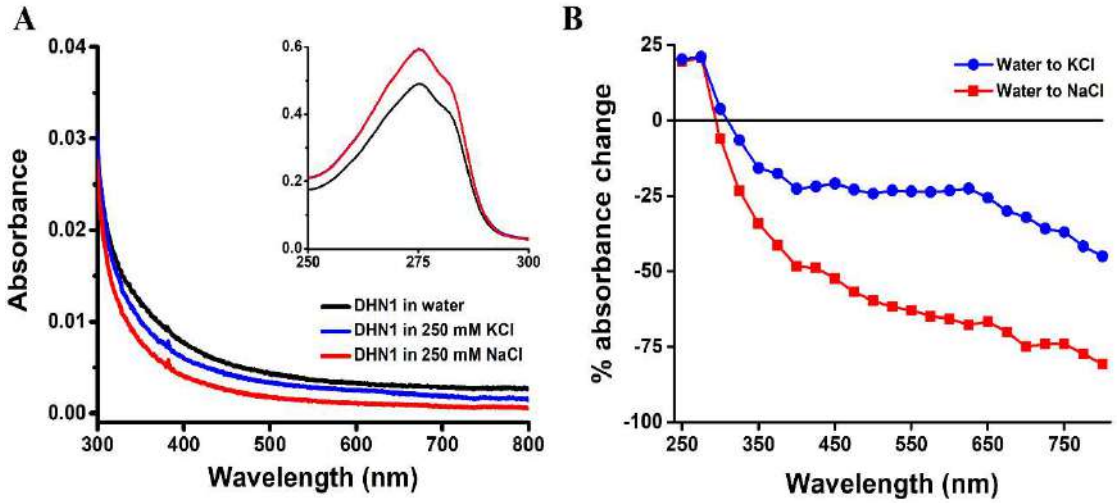


Figure 4.6 Effect of excess salt on ProCharTS: The effect of 250 mM KCl and 250 mM NaCl on absorption spectra of **(A)** DHN1 dissolved in water is shown. In inset spectrum shown for shorter wavelength (250-300 nm) of DHN1 in expanded form. **(B)** % Change in absorbance at selected wavelengths were calculated as [(Absorbance in salt – Absorbance in water)/ Absorbance in water] × 100.

TH-3004_146106002

4.3 Conclusion:

- Protein Charge Transfer Spectra (ProCharTS) was distinctly observed in DHN1 protein from 300- 800 nm wavelength range.
- Interference due to strong absorption arising from aromatic amino acid was absent in the 300-800 nm wavelength region, which assigns this as a distinct universal absorption band for ProCharTS.
- ProCharTS absorption was sensitive to changes in structure of DHN1 due to change in pH.
- An increase in ProCharTS intensity with rise in temperature correlated with change in the secondary structure of DHN1 protein as shown by CD.
- ProCharTS intensity was sensitive to the presence of salts in the medium.
- ProCharTS absorption can be used as a label-free intrinsic probe to monitor structural transitions in charged-rich proteins and IDPs.

TH-3004_146106002

Chapter 5

Luminescence characteristics of ProCharTS in DHN1

TH-3004_146106002

5.1 Introduction:

Label-free approaches to investigate the structure and track the subcellular location of biomacromolecules such as proteins are in demand. Such methods enable protein observations to be made in their natural state without any perturbation in their chemical or cellular composition[204]. Moreover, label-free approaches decrease the requirement of toxic dyes/probes, hence endorsing green chemistry[205]. In this regard, intrinsic luminescence could be used as a label-free tool to investigate proteins.

In naked proteins without cofactors like metal or coenzymes, the aromatic amino acid; tryptophan, tyrosine, and phenylalanine are believed to be the sole chromophore behind the origin of luminescence[156]. However, a protein class like intrinsically disordered proteins (IDPs) are mostly devoid of these chromophores in their sequence. Aside from these intrinsic chromophores, it is generally believed that monomeric proteins have no other sources of intrinsic luminescence.

A decade ago, our group reported the novel electronic absorption and subsequent luminescence originating from dense aqueous solutions of L-Lysine monohydrochloride[136]. Further, this novel finding was also supported by other research groups as well[135,138,139]. The maximum absorption of Lysine·HCl was observed at ~270 nm, and upon excitation at 355 nm, luminescence emission maxima was observed at ~435 nm. Similar observations have also been reported for the dilute solutions of Human Serum Albumin (HuSA) protein, which is rich in charged amino acids like Lys and Glu (Homchaudhuri and Swaminathan, unpublished). With the similar absorption spectra observed in L-Lysine solutions and protein rich in lysine amino acid residue, it was believed that the reason behind such novel absorption and luminescence could be inter/intra molecular interaction between the lysine residue. However, this observed luminescence was not apparent, as the lysine amino acid has an aliphatic side chain with an amine headgroup. To examine the cause behind this unusual observation, our group and collaborator in 2017 used a combination of experimental and theoretical approaches on a synthetic protein α_3 C rich in charged amino acid residues like Lys/Glu and devoid of aromatic amino acids. The findings suggest that the absorption spectra in the 250- 800 nm range could originate due to charge transfer (CT)

TH-3004_146106002

transitions involving the protein backbone and the charged residues in protein[144]. Similarly, our collaborator also showed using the time-dependent density functional theory (TDDFT) calculations, that Protein Charge Transfer Spectra (ProCharTS) arising from the other naturally occurring charged amino acids (Arg, Asp, His) and post-translational modifications like phosphorylation of Ser/Thr/Tyr[145].

Changes in ProCharTS absorption intensity upon changing pH, addition of salt, and increasing temperature alter the protein charge distribution, neutralize charges, and disrupt protein structure, respectively, were described in chapter 4. Plenty of work has been done so far[136,137,144,145,199], on the absorption arising from the charged amino acids. However, little work has been done so far upon luminescence arising from the ProCharTS. Earlier, the concentrated solutions of L-Lysine (0.5M) displayed luminescence upon excitation at different λ_{ex} 290-410 nm.

In this chapter, we looked into the photochemical feature of intrinsic luminescence arising from monomeric DHN1 protein from charge recombination after excitation at ProCharTS absorption bands. We studied the DHN1 protein luminescence properties, quantum yield, lifetime, Stokes shift, and excitation/emission spectra. Further, we investigated the effect of weak intrinsic luminescence on the tryptophan lifetime in DHN1 mutant; DHN1 CW1.

TH-3004_146106002

5.2 Results and Discussion:

5.2.1 Luminescence Spectra of DHN1:

In the previous chapter, we reported the ProCharTS absorption spectra of DHN1 protein and its sensitivity toward the protein charge distribution, neutralization of charge, and disruption of protein structure. Further, our group showed that lysine·HCl solutions shows luminescence upon excitation at different wavelengths. To further investigate whether ProCharTS absorption observed in charged abundant DHN1 protein is emissive or not, we monitored the emission spectra of DHN1 protein at different excitation wavelengths λ_{ex} 310 nm, 340 nm, 355 nm, 370 nm, and 410 nm. **Figure 5.1A** shows the emission spectra of DHN1 with increasing excitation wavelengths. Similarly, **Figure 5.1B** shows the normalized luminescence intensity of the DHN1, highlighting the consistent red shift in the emission peak with increasing excitation wavelength.

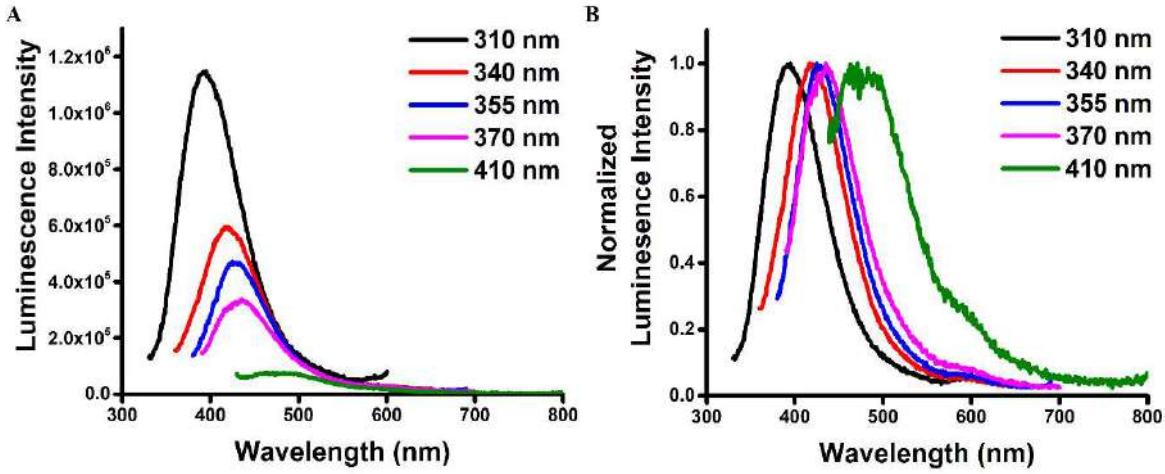


Figure 5.1. Luminescence emission of 60 µM of DHN1 (A) Luminescence spectra, and **(B)** Peak normalized luminescence spectra at different excitation wavelengths. Excitation and emission slit width was 2 and 15 nm, respectively at each excitation wavelength.

DHN1 shows the emission maxima λ_{em} at 395 nm, 417 nm, 425 nm, 435 nm, and 471 nm for excitation wavelength λ_{ex} 310 nm, 340 nm, 355 nm, 370 nm, and 410 nm, respectively. The consistent decrease in luminescence intensity of DHN1 protein was observed from λ_{ex} 310- 410 nm.

TH-3004_146106002

5.2.2 ProCharTS Luminescence Excitation Spectra:

The emission spectra displayed in **Figure 5.1 show** that the emission maxima are between 400 nm to 500 nm when the sample is excited at 340, 355, 370 or 410 nm. Therefore, to reveal the excitation bands, excitation spectra were measured at three different emission maxima wavelengths, i.e., 425 nm, 450 nm, and 480 nm. **Figure 5.2 A, B and C** displays the observed excitation spectra at λ_{em} fixed at 425 nm, 450 nm, and 480 nm, respectively.

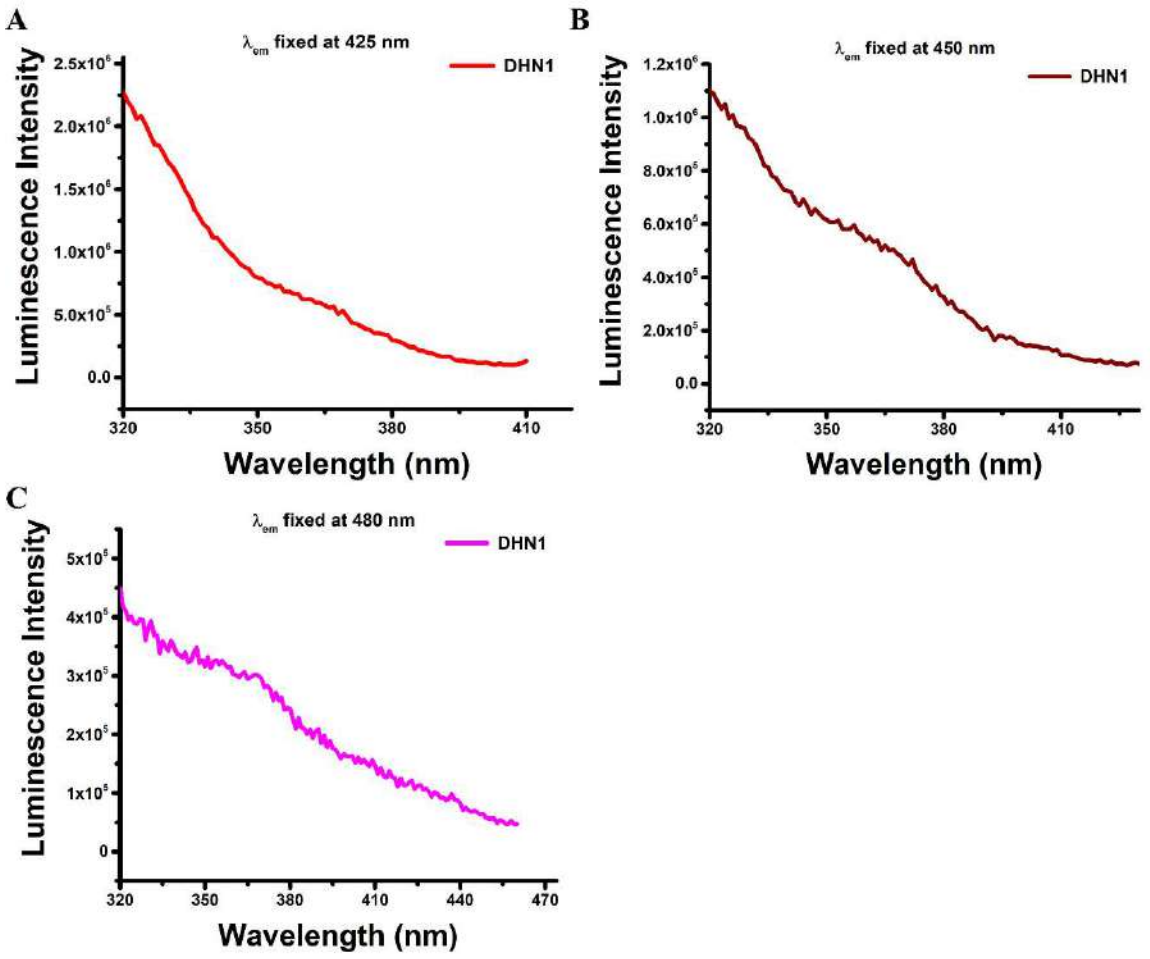


Figure 5.2. Luminescence excitation spectra of 60 µM of DHN1 at λ_{em} fixed at (A) 425 nm, **(B)** 450 nm, and **(C)** 480 nm.

When excitation spectra were measured at a fixed λ_{em} of 425 nm, the excitation bands for the DHN1 protein were found to be around 350 nm. While, for excitation measured at a fixed λ_{em} of 450 nm, the excitation band was found between 350-365 nm**.** Similarly, excitation bands were measured at a fixed λ_{em} of 480 nm, the excitation band was found

TH-3004_146106002

between 358 -380 nm. However, the excitation band is broadly distributed as observed in excitation spectra.

5.2.3 Luminescence linearity of DHN1:

To ascertain that ProCharTS absorption observed in DHN1 protein comes from the protein's monomeric state, the absorption spectra of DHN1 protein were measured as discussed in **section 4.2.1** of chapter 4. It was seen that the DHN1 showed a linear increase in absorption and thus suggested that ProCharTS absorption comes from the monomeric protein. Further to check the same in ProCharTS luminescence, we monitored the luminescence of DHN1 protein ranging from 10 to 60 µM at λ_{ex} of 355 nm. **Figure 5.3 A and B** shows the luminescence spectra and integrated luminescence area of DHN1 protein with increasing concentration at λ_{ex} 355 nm, respectively. DHN1 protein showed a linear relationship with an increase in protein concentration with an R^2 value of >0.99, as shown in **Figure 5.3B**. The slope from the linearity plot reveals the increase in luminescence intensity for unit protein concentration. The slope value obtained for the DHN1 protein is 1.27×10^7 (Intensity/ µM). This lower slope value compared with other folded proteins like α_3C and HuSA[146] was due to less effective concentration of charged amino acid clusters in DHN1 protein owing to its intrinsic disordered structure. Linearity in the ProCharTS luminescence

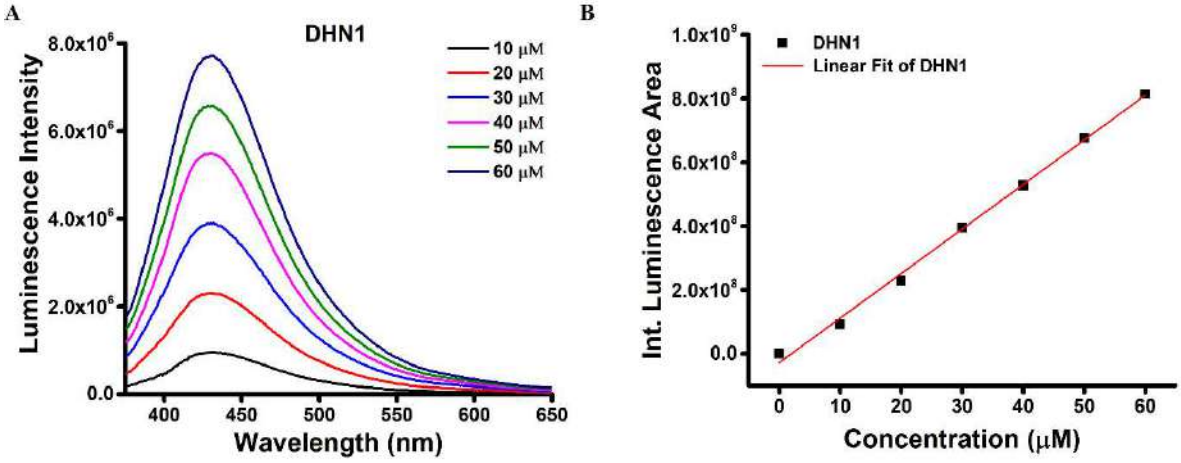


Figure 5.3. Luminescence emission of 60 µM of DHN1 at λ_{ex} 355 nm (A) Luminescence spectra, and **(B)** rise in integrated luminescence fitted linearly with increasing concentration are shown. All data was collected in duplicates and 10 scans were recorded for each concentration. Excitation and emission slit width was 2 and 15 nm, respectively. The λ_{em} maxima for DHN1 is at 425 nm.

TH-3004_146106002

of DHN1 protein is in close agreement with the ProCharTS absorption linearity observed previously, in chapter 4 **(Figure 4.2).** However, our results displayed the emission strictly coming from the monomeric state of protein in contrast to findings of other groups that reported that such emission may originate from the oligomeric state of proteins[148], crystal or protein aggregates[141,142].

5.2.4 Integrated luminescence at different wavelengths:

ProCharTS luminescence yield of DHN1 protein was calculated at different excitation wavelengths. **Figure 5.4** shows the integrated luminescence of DHN1 protein at different excitation wavelengths. The luminescence yield of DHN1 protein is maximum at 310 nm and further shows the decreasing trend from 310 to 410 nm. It is evident from the luminescence emission of DHN1 protein that the integrated luminescence intensity beyond 355 nm λ_{ex} is relatively low in comparison with excitation wavelengths <355 nm. Due to this reason, the linearity in luminescence emission and quantum yield for DHN1 proteins was measured at λ_{ex} of 355 nm. Further, with λ_{ex} of 355 nm, emission from Trp/Tyr was

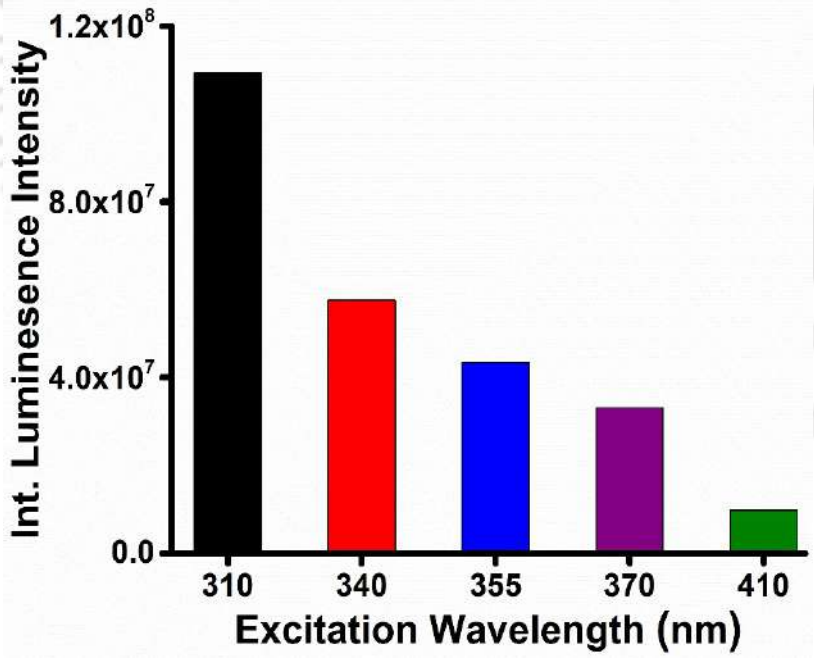


Figure 5.4. Integrated luminescence of 60 µM of DHN1 at different wavelengths. DHN1 protein was excited with 2 nm of slit width and emission was collected with 15 nm of slit width at the given excitation wavelength. Integrated luminescence was calculated for the emission spectral range containing >90% of luminescence for the given λ_{ex}.

TH-3004_146106002

also avoided. DHN1 protein shows the similar trend as observed for the HuSA, PEST, α_3C, and α-synuclein proteins and charged amino acids[146]. This result suggests a common phenomenon namely radiative charge recombination, responsible for the observed luminescence.

5.2.5 Stokes shift at different wavelengths:

Further, the Stokes shift as a function of excitation wavelength was calculated from the emission spectra of DHN1 protein. **Figure 5.5** shows the excitation wavelength-dependent Stokes shift observed for DHN1 protein. The decrease in Stokes shift from >6942 to >3159 cm^{-1} with the increase in excitation wavelengths (310 nm to 410 nm) is noticeable in **Figure 5.5**. However, the change in Stokes shifts with increasing excitation wavelength should be noted, which does not follow the Kasha's rule[206]. This could be attributed to the intramolecular charge transfer[207,208], and the formation of different charge transfer states[209]. Further, the large Stokes shift obtained for the DHN1 protein enriched with charged amino acids gives an impression of considerable change in the dipole moment in the excited state of these molecules. As Trp exposed to the solvent upon excitation at 280 nm display the emission peak ~348 nm with a Stokes shift of about ~7000 cm^{-1}, [183]. Here the charged amino

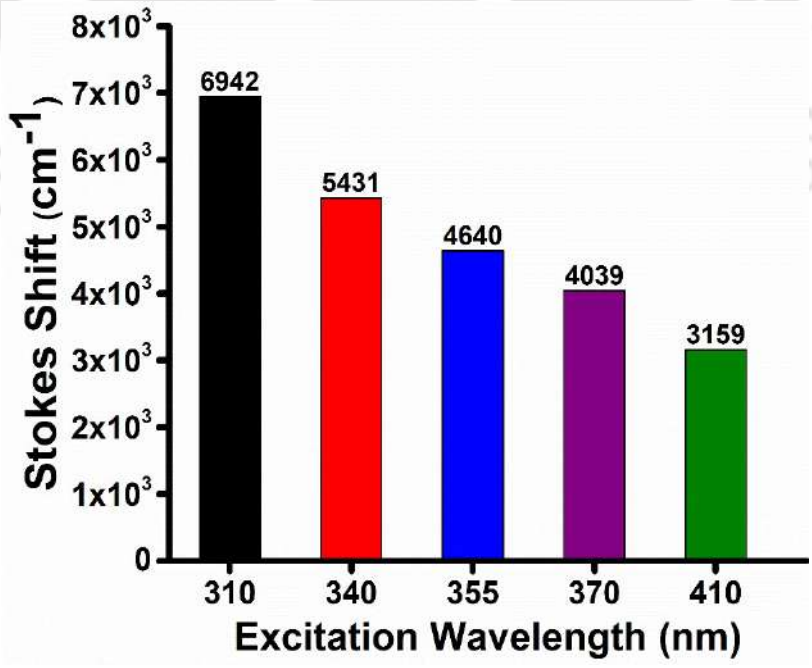


Figure 5.5. Stokes shift of 60 µM of DHN1 as function of excitation wavelength.

TH-3004_146106002

acid enriched DHN1 protein shows the huge Stokes shift at wavelength beyond 310 nm. This suggests the higher sensitivity of luminescence arising from the charge transfer state towards the polarity of the solvent.

Further, the quantum yield of DHN1 was calculated at λ_{ex} 355 nm, as defined in the methods (chapter 2). The calculation of quantum yield at λ_{ex} 355 nm is done for two reasons. First, this wavelength (355 nm) lies within the range of excitation bands for DHN1. Second, any contribution from aromatic amino acids is easily avoided (if present). The calculated quantum yield for DHN1 was 0.027. The absorbance and emission spectra used for the calculation of the quantum yield of DHN1 are shown in the appendix (**Figure A2)**.

The observed quantum yield is significantly lower than the conventional fluorophores viz., tryptophan, present in protein, which is 0.14,[210,211]. DHN1 protein shows the similar shape and Stokes shifts in the emission spectra as reported for other folded proteins and IDPs like HuSA, PEST, α_3 C, and α-synuclein[146]. These results suggest that the DHN1 has the same photophysical mechanism behind the observed luminescence.

In next section, we monitored the luminescence lifetime of DHN1 protein at 295 and 340 nm to further characterize the ProCharTS luminescence properties.

5.2.6 Luminescence lifetime of DHN1:

To characterize the time-resolved luminescence properties of DHN1, luminescence intensity decay at λ_{ex} 295 and 340 nm was collected. The λ_{ex} 295 nm was used to compare with the traditional chromophores, Trp from the same spectral region (250-300 nm). While the λ_{ex} 340 nm wavelength was selected as it was within the range of excitation bands for DHN1 protein. **Figure 5.6 A and B** show the luminescence intensity decay along with the fit of DHN1 protein at $\lambda_{ex\ of}$ 295 and 340 nm, respectively. DHN1 protein shows the luminescence intensity decay that fit best to three distinct lifetimes with changing amplitudes. **Figure 5.7** shows the trend in lifetime values at λ_{ex} 295 nm and 340 nm for DHN1 protein. For λ_{ex} of 340 nm, that display a very short lifetime of ~1 ns followed by two longer lifetimes ~3 ns and ~10ns. The shortest component (τ_3) has maximum amplitude values ~0.8. Similarly, at λ_{ex} 295 nm DHN1 shows the three different lifetimes with different amplitudes. In this case, the smallest component (τ_1) is <1 ns, while the other two

TH-3004_146106002

components have values of ~3 ns (τ_2) is >0.5, and the longest component (τ_3) shows the contribution >0.1. The parameter obtained from the discrete fit are given in the **Table 5.1**.

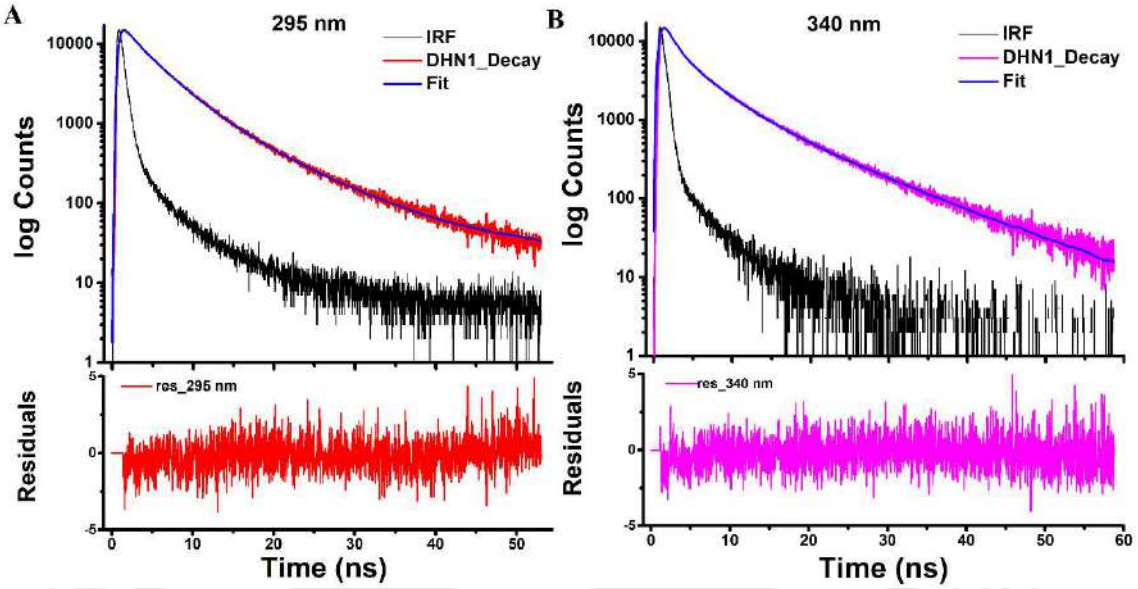


Figure 5.6. Luminescence intensity decay profile of 60 µM of DHN1. Three exponential intensity decay fit for DHN1 at λ_{ex} at **(A)** 295 nm, and **(B)** 340 nm. Residuals for fit is shown below the decay fit. Emission was collected for λ_{ex} 295 and 340 nm using 320 nm and 370 nm long pass filter, respectively. All the decay was collected in duplicates and best fitted data is shown here.

Table 5.1. Parameters from three exponential fit for luminescence intensity decay of DHN1 obtained for λ_{ex} at 295 and 340 nm. Values in the brackets (for τ_i and τ_{mean}) indicate the standard deviation observed in 3 independent measurements, while the α_i are represented from the best fit.

DHN1 λ_{ex} **(nm)**	**τ_1 (ns)**	**τ_2 (ns)**	**τ_3 (ns)**	**α_1**	**α_2**	**α_3**	**τ_m (ns)**	**χ^2_R**
295	0.64 (0.10)	3.4 (0.01)	8.3 (0.03)	0.35	0.54	0.111	2.9. (0.13)	1.11
340	1.02 (0.03)	3.9 (0.11)	10.6 (0.18)	0.59	0.36	0.088	2.8 (0.04)	1.2

TH-3004_146106002

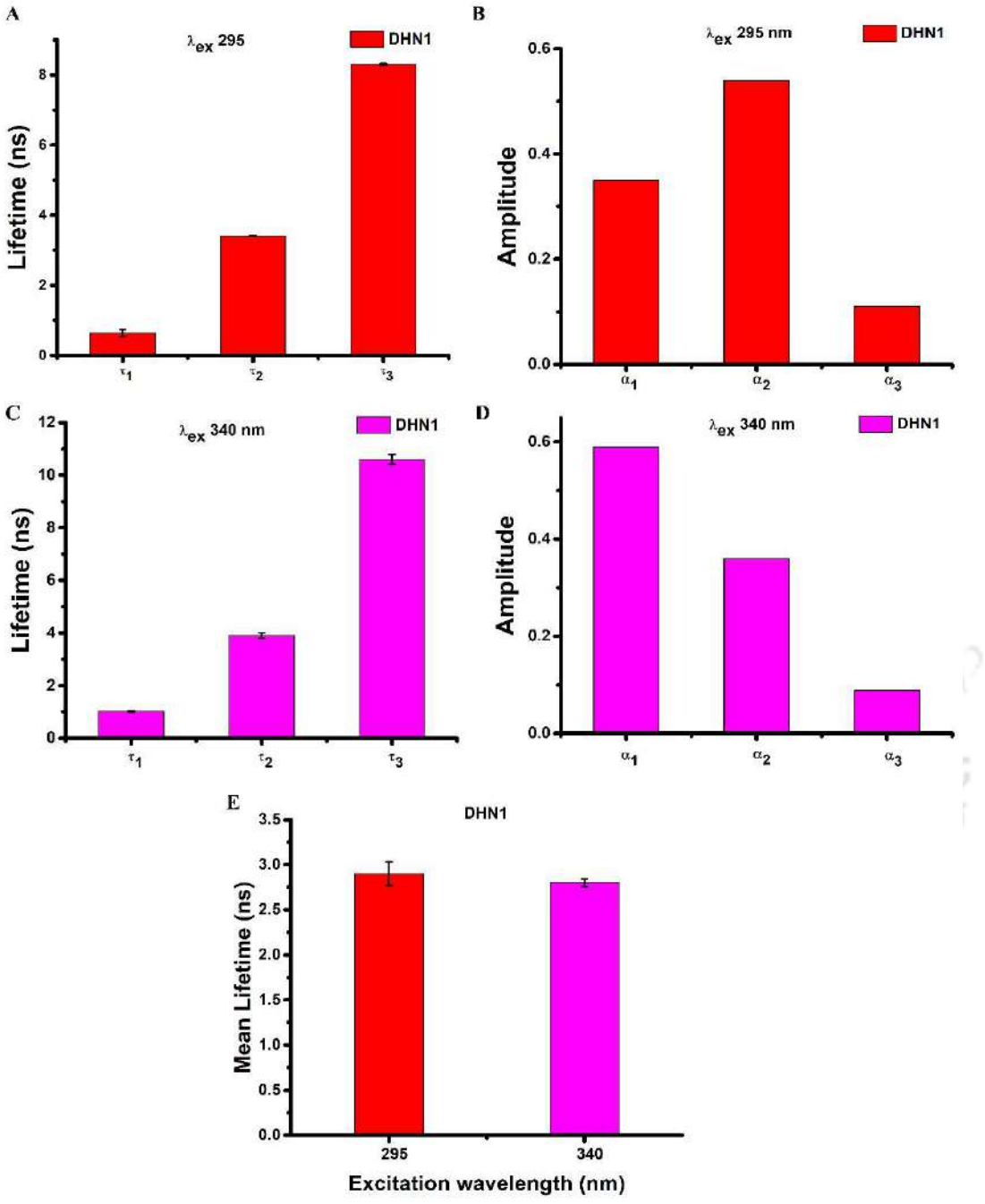


Figure 5.7. Discrete analysis of ProCharTS luminescence intensity decay of DHN1 at λ_{ex} 295 and 340 nm. **(A)** Lifetime values at λ_{ex} 295 nm, **(B)** Amplitude at λ_{ex} 295 nm; **(C)** Lifetime values at λ_{ex} 340 nm; **(D)** Amplitude at λ_{ex} 340 nm; and **(E)** mean lifetime values at λ_{ex} 295 and 340 nm are shown.

TH-3004_146106002

The decay profile of DHN1 luminescence at λ_{ex} 295 and 340 nm clearly reveals the multi-exponential decay. The mean lifetime which represents the area under the luminescence decay curve is proportional to the steady state luminescence intensity of the protein. The mean lifetime of DHN1 obtained from the discrete analysis at λ_{ex} 340 and 295 nm was 2.8±0.1 ns and 2.9±0.1 ns, respectively, as shown in **Figure 5.7; Table 5.1**. This higher mean lifetime in DHN1 is consistent with the high quantum yield when compared to proteins like HuSA, PEST, α_3 C, and α-synuclein[146].

This discrete analysis of luminescence intensity decay was further validated with the Maximum Entropy Method (MEM) analysis (explained in Chapter 2).

For λ_{ex} 295 nm, DHN1 protein shows the three luminescence lifetime distributions, i.e., three lifetimes with varying amplitude, as displayed in **Figure 5.8A.** The shortest component observed for DHN1 is >0.2 ns. The second component with a peak value of 3.1 ns, and the third one with a peak value around ~10 ns were observed. Similarly, for λ_{ex} 340 nm, the three peak distributions corresponding to three different lifetimes were observed. **Figure 5.8B** show the lifetime distribution of DHN1 at λ_{ex} 340 nm. The shortest component observed for DHN1 protein is ~0.2 ns. The longest component observed for the DHN1 is around 8-9 ns. The residuals to evaluate the goodness of fit for MEM analysis is shown by the best residual fit below the MEM distribution curve for λ_{ex} 295 nm and 340 nm in **Figure 5.8.**

The MEM result reveals that the shortest component showed the very broad distribution and maximum amplitude than the other two components, while the longest component has shown the least contribution (in case of λ_{ex} 340 nm)**.** Whereas in λ_{ex} 295 nm, the second component showed a maximum amplitude value than the rest of the other two components. The shortest component displayed a broad distribution similar to λ_{ex} 340 nm. With the longer lifetime values for the first and second components and with the highest contribution from the second component, DHN1 can be predicted to have the longest mean lifetime (in comparison with HuSA, PEST, α_3C and α-synuclein proteins[146]) consistent with the longer mean lifetime observed from the discrete analysis. Similarly, the highest contribution from the second component (3.1 ns) is also comparable to the results of discrete analysis, where the contribution (α_2) from the second component (τ_2) is the maximum.

TH-3004_146106002

The similar lifetime distribution obtained for λ_{ex} 340 nm and 295 nm in DHN1 protein suggests similar charge transfer transitions occurring at different wavelengths.

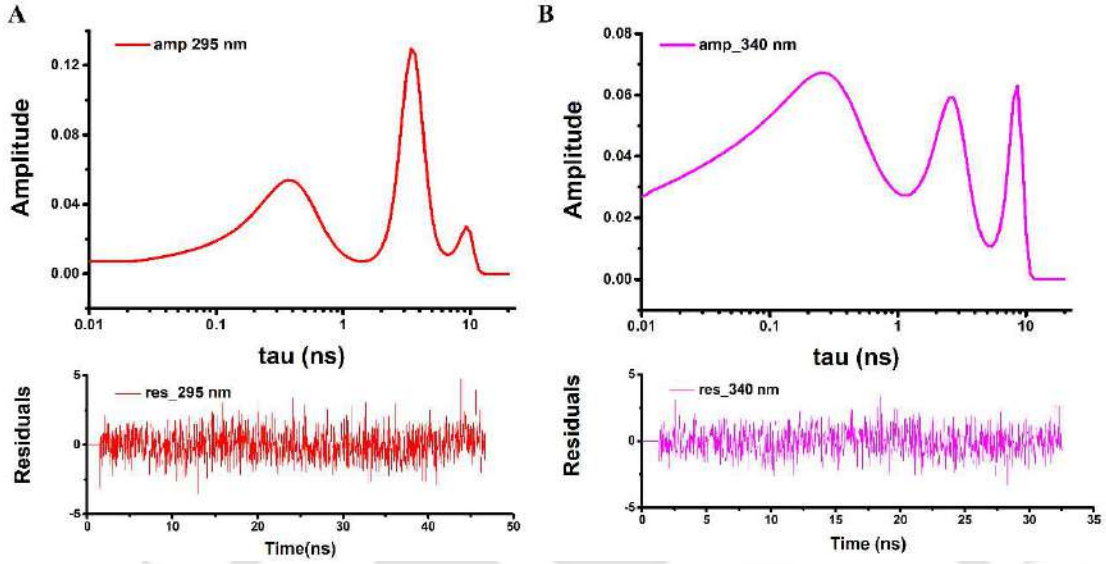


Figure 5.8. Analysis by Maximum Entropy Method. For MEM analysis, luminescence decay of each protein was fitted in 100 exponentials in the range of 0.01-20 ns with 0.001 as the initial amplitude value for the distribution. Panel shows the lifetime distribution of DHN1 for **(A)** λ_{ex} done at 295 nm, and **(B)** λ_{ex} done at 340 nm.

Further, to ascertain the role of excitation pulse width and emission wavelength, luminescence intensity decay of DHN1 protein were measured using a temporary shorter excitation pulse (~80 ps IRF) along with emission monochromator or band pass filter, as shown in **Figure 5.9A.** The DHN1 protein was excited at λ_{ex} 307 nm, and 369 nm and the emission were collected at λ_{em} 439 nm, and 437 nm, respectively. The residuals for discrete fit are shown in **panel A of Figure 5.9**. The discrete analysis of DHN1 luminescence intensity decay reveals the three distinct lifetimes of DHN1; parameters and lifetime values obtained from the discrete fit are shown in **Table 5.2**.

The normalized lifetime distribution obtained from the MEM analysis of DHN1 luminescence intensity decay and residuals for the MEM fit is illustrated in **Figure 5.9C.** The lifetime distribution obtained at shorter excitation pulse also showed the similar

TH-3004_146106002

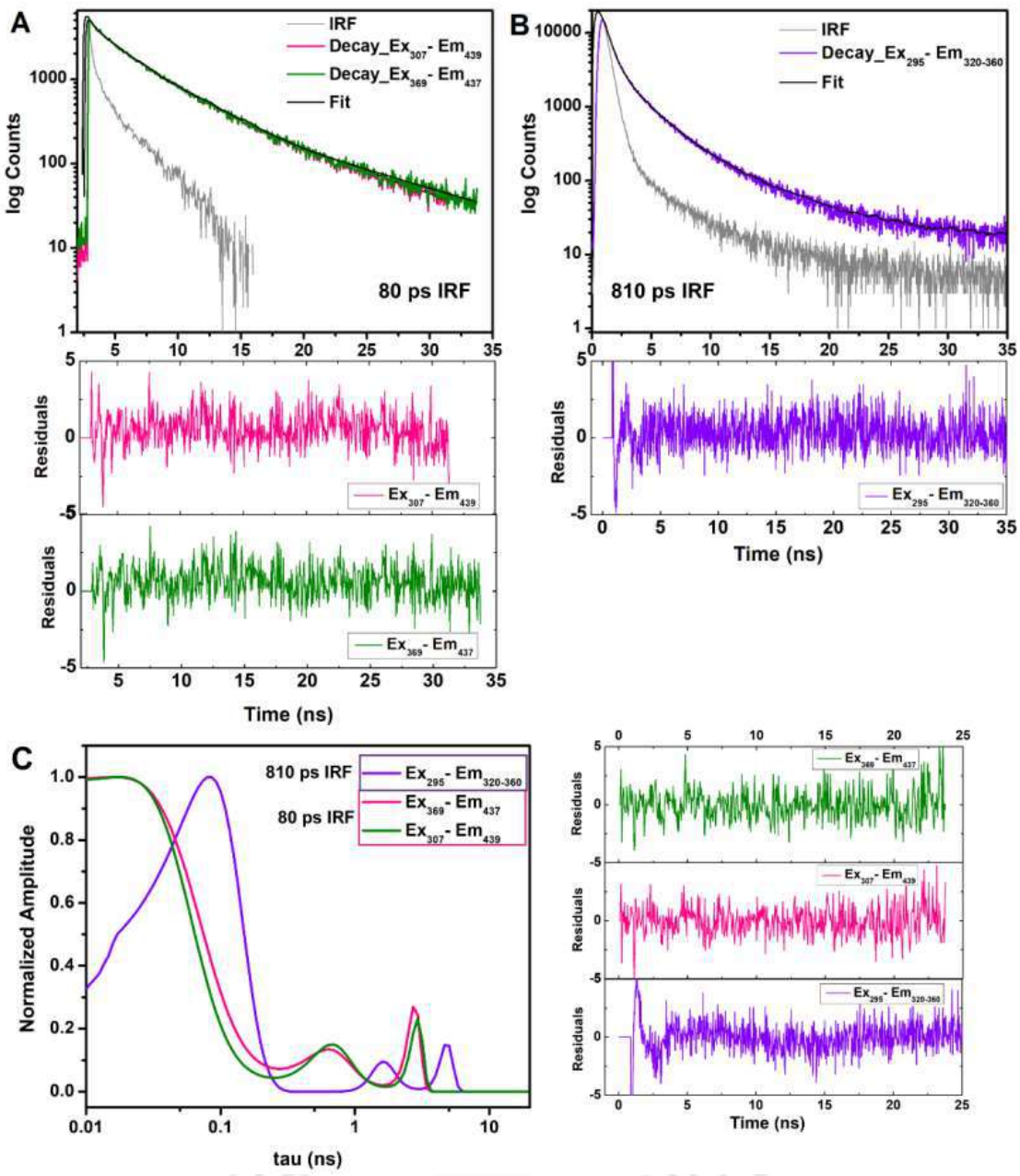


Figure 5.9. Comparative analysis of luminescence decay of DHN1 using laser source with shorter pulse width: Panel **A** displays the luminescence intensity decay (with their discrete fit) at two different excitation wavelengths (307 nm and 369 nm) with emission collected at specific wavelengths (439 nm and 437 nm, respectively) The data were collected with very short pulse width (80 ps) laser (see details below). Panel **B** displays the luminescence intensity decay collected for excitation at 295 nm and emission collected with bandpass filter (340 ±20 nm). This data was collected using similar laser with higher pulse width (810 ps) as described in section **2.3.4 (chapter 2)**. Panel **C** displays the lifetime distribution obtained from MEM analysis for the identical samples. Residuals from the respective fit are displayed in lower subpanels for A and B in adjacent subpanel for C.

TH-3004_146106002

three peaks corresponding to three different lifetimes with varying amplitudes, as shown previously in **Figure 5.8 A, and B**.

The collection of emission photons in a restricted wavelength band seems to reduce the mean lifetime of DHN1, as shown in **Table 5.2.** However, using a shorter excitation pulse does not significantly resolve the broad lifetime distribution in the 0.01-0.1 ns range as the time resolution of data acquisition (48ps/channel) is not better than 28 ns/channel used in this work. Perhaps the fluorescence upconversion technique is required to resolve the shorter lifetimes observed.

Table 5.2. Parameters from three exponential fit for luminescence intensity decay of DHN1 obtained with different pulse width lasers. Values in the brackets (for τ_i and τ_{mean}) indicate the standard deviation observed in 3 independent measurements, while the α_i are represented from the best fit.

DHN1			τ_1	τ_2	τ_3	α_1	α_2	α_3	τ_m	χ_R^2
λ_{ex} (nm)	λ_{em} (nm)	IRF (ps)	(ns)	(ns)	(ns)				(ns)	
307	**439**	**80**	0.19 (0.01)	1.42 (0.42)	7.83 (0.16)	0.816	0.155	0.029	0.63 (0.01)	1.53
369	**437**	**80**	0.22 (0.01)	1.76 (0.05)	9.98 (0.27)	0.808	0.020	0.020	0.66 (0.04)	1.5
295	**320-360**	**810**	0.39 (0.01)	1.93 (0.05)	4.84 (0.05)	0.905	0.011	0.011	0.57 (0.01)	1.34

5.2.7 ProCharTS luminescence can influence fluorescence decay of tryptophan in protein:

With the broader absorption range of ProCharTS across the UV-Visible spectrum[144–146,199], it is interesting to divulge its effect on several intrinsic and extrinsic probes. The intrinsic fluorophores in the side chain of aromatic amino acids like Trp, Tyr, and Phe are believed

TH-3004_146106002

to be the sole origin of intrinsic fluorescence among the proteins. Moreover, the indole ring in the tryptophan is frequently used as a fluorescent probe in proteins. However, our study on ProCharTS absorption and luminescence spectra revealed that the prevalence of ProCharTS in the broad spectral range (250-800 nm) could potentially influence the fluorescence signal from the tryptophan in the proteins. To investigate the effect of this hidden unconventional chromophore on the tryptophan emission, we studied the NATA (a Trp analogue) fluorescence lifetime in presence of charged-rich DHN1 protein. Subsequently, the comparative study of DHN1 and its tryptophan mutant DHN1 CW1 was also carried out.

5.2.7.1 Deviation from single exponential fluorescence decay of NATA in presence of Lysine-rich DHN1 protein

As previously explained in this work, the charged and non-aromatic amino acid residues exhibit considerable absorption and luminescence in the same spectral region where only aromatic amino acids were supposed to show absorption and fluorescence. The possibility that it could interfere with the intrinsic and extrinsic fluorescence probe can not be neglected. To investigate whether the charged and non-aromatic amino acids residues could influence the fluorescence decay of tryptophan in proteins, we have performed time-resolved fluorescence studies of N-acetyl-L-tryptophan amid (NATA) which has single exponential decay at λ_{ex} 295 nm, in presence of an increasing concentration of DHN1. Here we exploited the amino acid composition of DHN1, which is devoid of Trp and enriched with 29 % of charged amino acid residues.

To investigate whether the NATA could bind non-specifically to the DHN1 protein, the steady state anisotropy measurement of NATA in the presence of DHN1 protein was performed. **Figure 5.10** displays the steady-state anisotropy of 20 µM NATA excited at λ_{ex} 295 nm (λ_{em} 350 nm) in the presence of an increasing concentration of DHN1 protein (10-100 µM). An increase in the steady-state anisotropy of NATA with the rise in the DHN1 concentration was observed. This result suggests that the NATA does bind to the DHN1 protein in a concentration-dependent manner. This gives a clue that the NATA is in close proximity to the DHN1 protein surface.

TH-3004_146106002

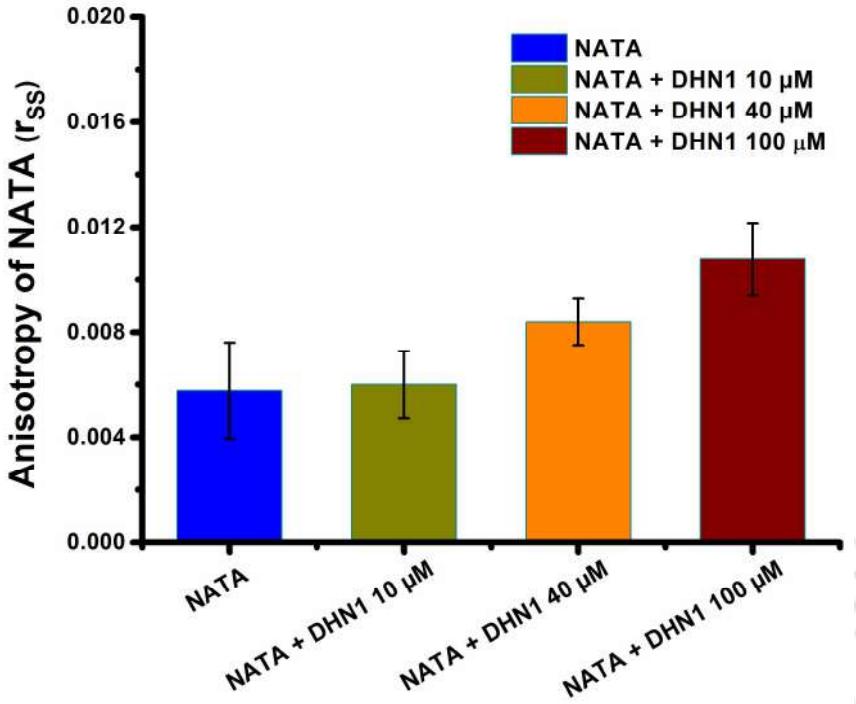


Figure 5.10. Steady-state anisotropy of 20 μM NATA (λ_{ex} 295 nm; λ_{em} 350 nm) in presence different concentration of DHN1.

Further, to understand whether the charged amino acid enriched-DHN1 protein can influence the fluorescence decay kinetics of NATA, time-resolved fluorescence intensity decay of NATA in the presence DHN1 protein (5-100 μM) was collected. **Figure 5.11A and B** display the time-resolved fluorescence intensity decay and single exponential fit for 20 μM NATA (λ_{ex} 295 nm), respectively, in the presence of different concentrations of DHN1 protein (5-100 μM). The parameters obtained from a single exponential discrete fit of time-resolved fluorescence intensity decay of NATA in the presence of DHN1 are shown in **Table 5.3**

TH-3004_146106002

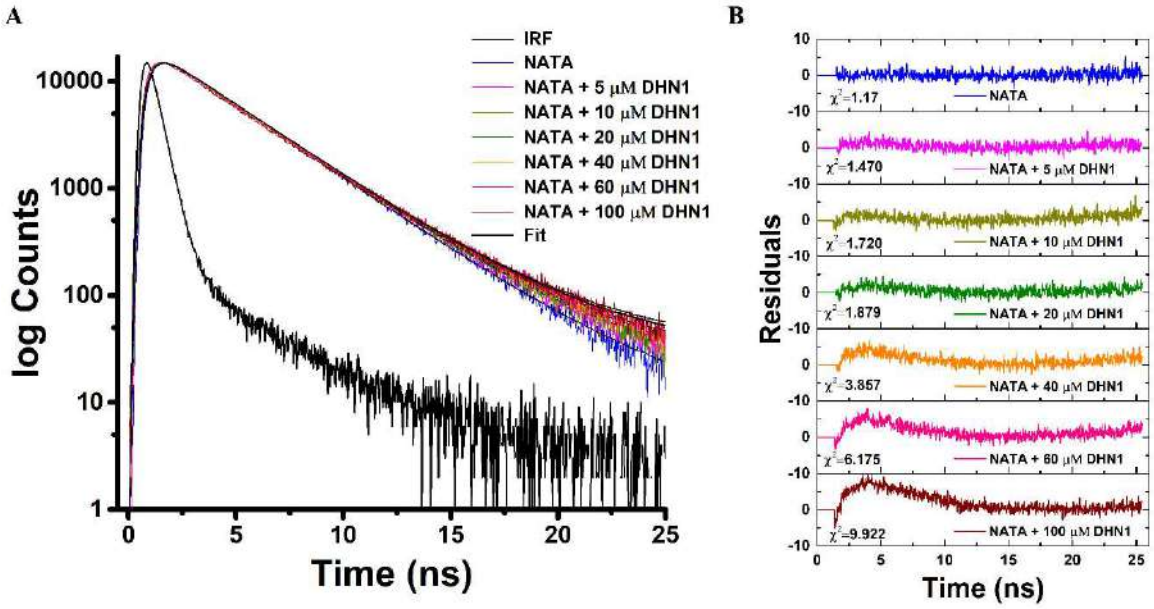


Figure 5.11. Time-resolved fluorescence intensity decay of NATA (20 μM) in presence of different concentration of DHN1. (A) Single exponential fit for NATA at λ_{ex} 295 nm, **(B)** Residuals for the fit. Emission was collected using 340± 20 nm band pass filter. All the decay was collected in duplicates and best fitted data is shown here.

Table 5.3. Parameters obtained from one exponential discrete fit on fluorescence intensity decay of NATA in the presence of DHN1. Values in the brackets (for τ_i and τ_{mean}) indicate the standard deviation observed in 3 independent measurements, while the α_i are represented from the best fit.

Sample	α_1	τ_m (ns)	χ^2_R
NATA	1	2.97 (0.04)	1.01
NATA + 5 μM DHN1	1	3.03 (0)	1.47
NATA + 10 μM DHN1	1	3.00 (0)	1.72
NATA + 20 μM DHN1	1	3.00 (0)	1.69
NATA + 40 μM DHN1	1	3.04 (0)	3.86
NATA + 60 μM DHN1	1	3.03 (0)	5.84
NATA + 100 μM DHN1	1	3.05 (0)	9.82

TH-3004_146106002

The discrete analysis of fluorescence intensity decay of NATA in presence of DHN1 indicates that the DHN1 causes the deviation from single exponential kinetics. Increased χ_R^2 values with increasing DHN1 concentration (**in Table 5.3)** and bad residuals **(in Figure 5.11B)** show that the decay is unsuitable for the single exponential model. Then, same decay when fitted with 2 exponential model as shown in **Figure 5.12 A**, it gave χ_R^2 values closer to unity (**in Table 5.4)** and better residuals (**Figure 5.12B**). **Figure 5.13** shows the trends χ_R^2 values obtained from fitting (1 and 2 exponential model) of fluorescence intensity decay of NATA with increasing DHN1 protein concentration. The same trend was also observed in the pure charged amino acids and other charged-rich proteins like HuSA, PEST, α_3C and α-synuclein[146]. This discrete data analysis suggests that the high content of charged amino acids like lysine in DHN1 protein influences the NATA fluorescence decay kinetics.

Further, the fluorescence decay kinetics of NATA were analyzed using the Maximum Entropy Method (MEM). The lifetime distribution obtained from the MEM analysis for NATA in presence of DHN1 protein displayed a decrease in the amplitude of the long lifetime component with an increase in protein concentration, as displayed in **Figure 5.14 A**. A shorter component (≤0.5 ns) has been observed in presence of DHN1 protein, which

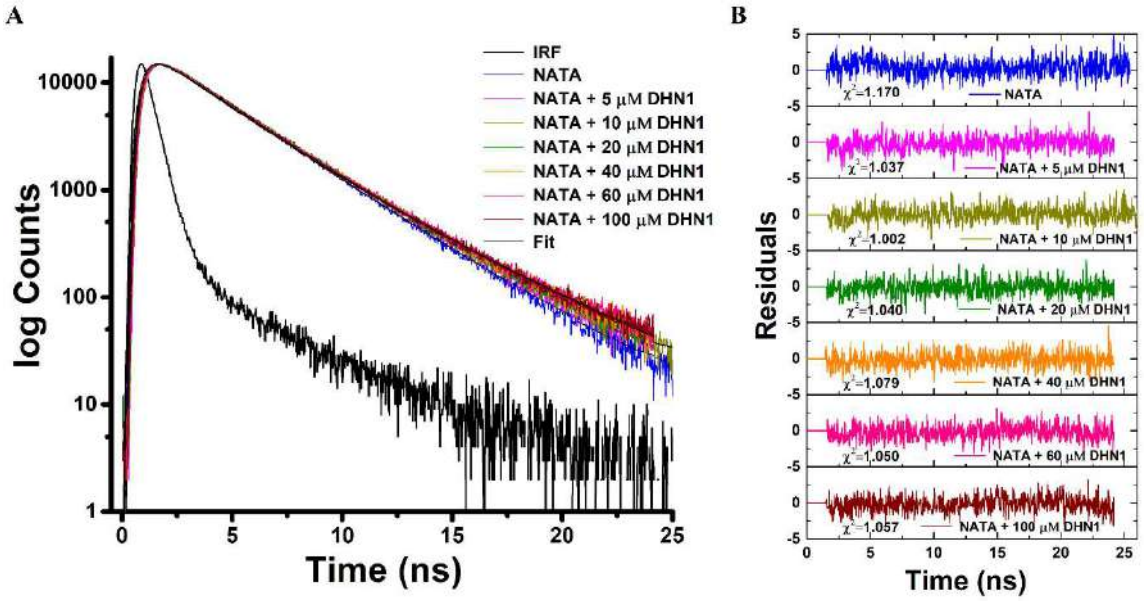


Figure 5.12. Time-resolved fluorescence intensity decay of NATA (20 μM) in presence of different concentration of DHN1. (A) Two exponential fit for NATA at λ_{ex} 295 nm, **(B)** Residuals for the fit. Emission was collected using 340± 20 nm band pass filter. All the decay was collected in duplicates and best fitted data is shown here.

TH-3004_146106002

could not be resolved from the discrete analysis (**Table 5.4**). This component was absent in lysine devoid PEST wt, His, Glu, Asp samples[146]. These observations hint at the role of charged lysine headgroup in influencing the fluorescence emission in NATA.

Table 5.4. Parameters obtained from two exponential discrete fit on fluorescence intensity decay of NATA (20 μM) in presence of DHN1. Values in the brackets (for τ_i and τ_{mean}) indicate the standard deviation observed in 3 independent measurements, while the α_i are represented from the best fit. NATA is fitted in single exponential model.

Sample	α_1	α_2	τ_1 (ns)	τ_2 (ns)	τ_{mean} (ns)	χ_R^2
NATA	1	--	2.97 (0.04)	--	--	1.01
NATA + 5 μM DHN1	0.936	0.064	2.97 (0.01)	4.35 (0.70)	3.04 (0)	1.04
NATA + 10 μM DHN1	0.192	0.808	2.31 (0.02)	3.18 (0)	3.01 (0)	1.00
NATA + 20 μM DHN1	0.151	0.85	2.29 (0.10)	3.16 (0)	3.02 (0)	1.04
NATA + 40 μM DHN1	0.771	0.229	2.74 (0)	3.93 (0.01)	3.01 (0)	1.08
NATA + 60 μM DHN1	0.732	0.269	2.68 (0)	3.87 (0)	3.00 (0)	1.05
NATA + 100 μM DHN1	0.661	0.34	2.50 (0.03)	3.86 (0.10)	2.96 (0)	1.04

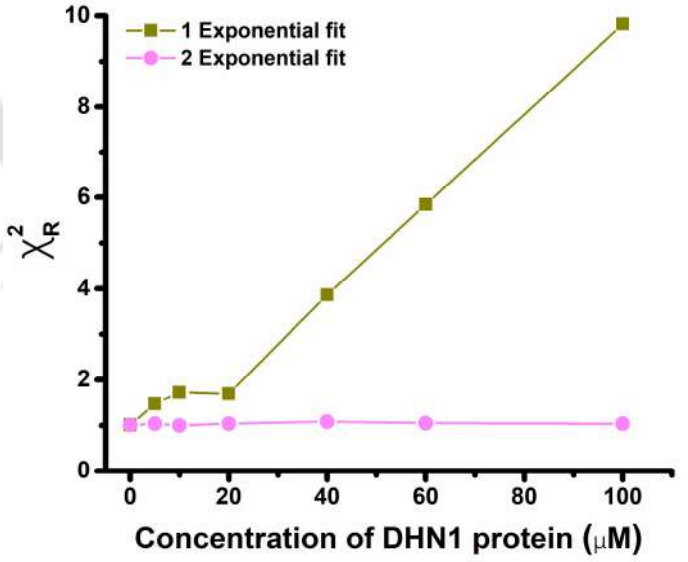


Figure 5.13. Values of χ_R^2 are plotted against the concentration of DHN1 protein in which decay of NATA (λ_{ex} 295 nm) was observed and fitted in either mono or bi-exponential model.

TH-3004_146106002

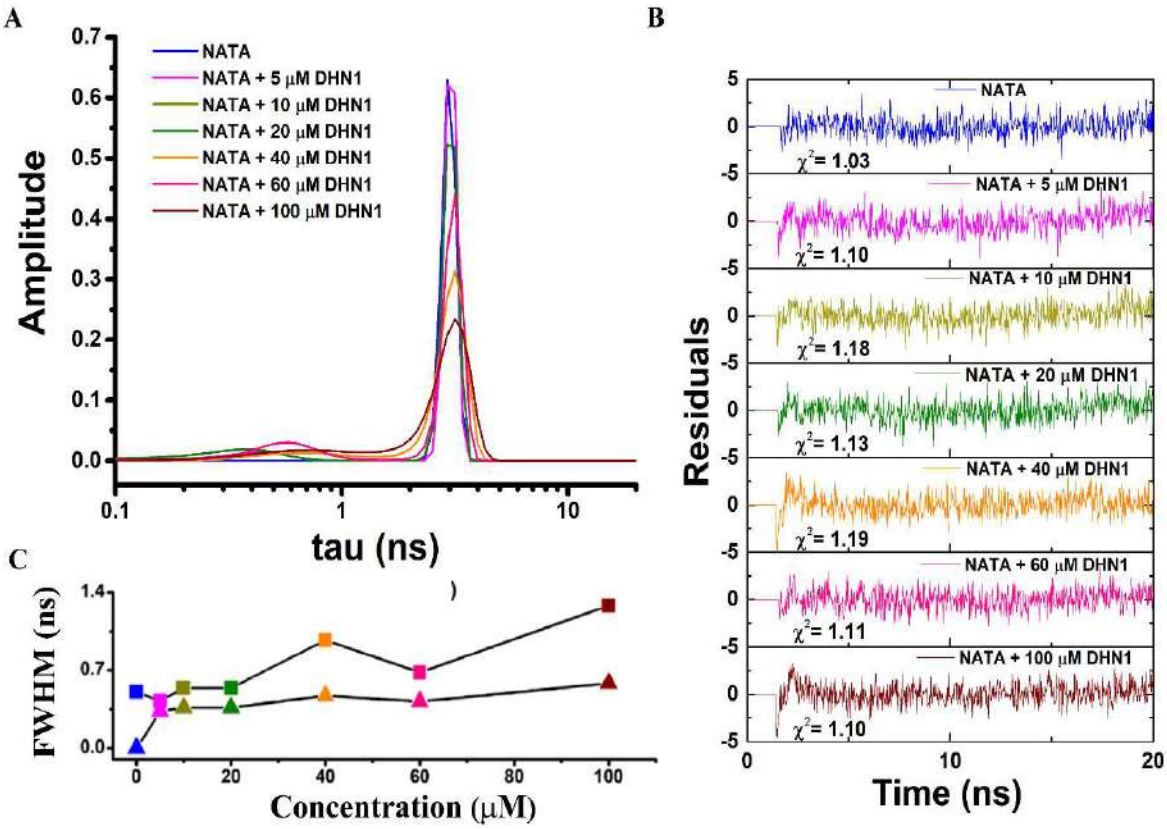


Figure 5.14. Maximum Entropy Method analysis. (A) Fluorescence lifetime distribution of NATA (λ_{ex} 295 nm) obtained using MEM, in presence of DHN1 protein at increasing concentrations (5-100 μM). **(B)** residuals for MEM fit; **(C)** the change in FWHM of the long(square) and short (triangle) lifetime components is shown.

The width of both the long and short lifetime components (higher FWHM values) displayed in **Figure 5.14C** shows an apparent increase in FWHM with an increased DHN1 protein concentration. Such an increase implies a broader distribution of lifetimes indicating heterogeneity in emission.

5.2.7.2 Fluorescence of Charge rich protein DHN1 in presence and absence of Tryptophan:

Next, we studied the intrinsic luminescence of DHN1 protein rich in charged amino acids after addition of tryptophan in the protein sequence (substitution mutation; DHN1 CW1). We look at the effects of ProCharTS on the steady-state emission spectra of Trp. **Figure**

TH-3004_146106002

5.15 A shows the peak intensity normalized fluorescence emission spectrum of NATA and DHN1 CW1 protein after Trp excitation at 295 nm. Compared to NATA control, which has no charged amino acids in the vicinity, DHN1 CW1 protein displays a marginally quenched emission, specifically on the longer wavelength side of the peak. This suggests a higher share of emission from ProCharTS bands than indole as we move towards longer wavelengths. Consistent with this possibility, we observe a small but significant increase in emission intensity after 450 nm in HuSA, PEST M1, once the Trp emission is adequately diminished[146]. This increase in emission on the red side also contributes to broadening of the Trp emission spectrum. The above result clearly shows that charged residues can alter

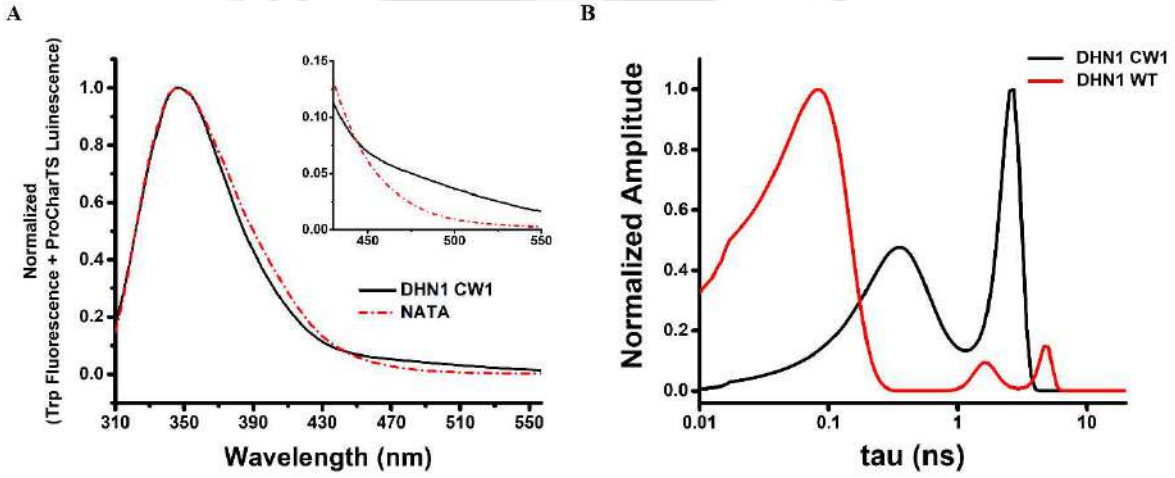


Figure 5.15. Fluorescence emission spectra; and lifetime distributions obtained from MEM for DHN1 and DHN1 CW1. (A) Peak Normalized emission spectra of DHN1 CW1 and NATA. The emission maxima for DHN1 CW1 and NATA are at 346 nm. Inset shows the enhanced emission in charge rich-DHN1 CW1 protein arising from ProCharTS luminescence in highlighted beyond the 450 nm. **(B)** Peak amplitude normalized fluorescence distribution of DHN1 and its Trp mutant at λ_{ex} 295 nm are shown.

the Trp emission spectrum both in terms of intensity and spectral distribution.

Next, we investigated the time-resolved fluorescence intensity decay of DHN1 and its mutant DHN1 CW1 protein **(shown in appendix Figure A3)**. The decay of the wild type DHN1 protein was fitted to three exponential model while the mutant DHN1 CW1 was fitted to two exponential model. The three different lifetimes observed for DHN1 protein are shown in **Table 5.5**. The short lifetime is ~0.39 ns with the maximum amplitude of ~0.9, and the other two are ~1.93 ns and ~4.82 ns with amplitude values of ~0.08 and ~0.01, respectively. The wild-type protein DHN1 shows the fast component (τ_1) ~0.39 with the

TH-3004_146106002

maximum contribution. Similarly, DHN1 CW1 protein shows two lifetimes ~2.67 (τ_1) and ~1.04 (τ_2) with the maximum amplitude contributed by $\tau_1 \geq 0.8$. The mean lifetime values observed for the DHN1 and its mutant upon excitation at λ_{ex} 295 nm DHN1 CW1 are 0.57 ns and 1.85ns, respectively.

Table 5.5 Parameters obtained from two or three exponential fit for luminescence intensity decay of DHN1 and its mutant DHN1 CW1 at λ_{ex} 295 nm. Emission was collected using the bandpass filter (340 ±20 nm). The Values in brackets (for τ_i and τ_{mean}) represents the standard deviation observed in two or more independent experiments while the α_i values are from the best fit.

Sample λ_{ex} 295 (nm)	τ_1 (ns)	τ_2 (ns)	τ_3 (ns)	α_1	α_2	α_3	τ_{mean} (ns)	χ_R^2
DHN1	0.39 (0.01)	1.93 (0.01)	4.82 (0.07)	0.90	0.08	0.01	0.57 (0.01)	1.11
DHN1 CW1	2.67 (0.03)	1.04 (0.05)	---	0.808	0.020	---	1.85 (0.02)	1.2

Figure 5.15 B displays the fluorescence lifetime distribution of DHN1 and its mutant DHN1 CW1 protein. The fluorescence lifetime distribution obtained with MEM analysis for DHN1 and its mutants was similar as observed in discrete analysis (**Table 5.5**). The fluorescence decay of wild-type DHN1 shows a fast decay component (10-100 ps) compared to its Trp mutant, hinting towards the much shorter lifetime (decay is not shown). This is further confirmed with the appearance of the short component between 0.2.-0.4 ns dominating the lifetime with maximum contribution of ≥0.9 (**Table 5.5**).

The time-resolved decay data indicate that when emission of Trp is collected at peak emission wavelength (350 nm), ProCharTS contribution is negligible. From the above illustrated multiple observation, it can be inferred that the existence of charge residues in the close proximity of Trp could affect its fluorescence lifetime and lead to multi-exponential decay kinetics of Trp proteins. The fluorescence intensity decay of the Trp explicitly depends on the nature of charge residues in the vicinity of the indole fluorophore.

TH-3004_146106002

5.3 Conclusion:

- An IDP abundant with charged residues, DHN1 is luminescent in monomeric states.
- The ProCharTS luminescence yield of DHN1 protein decrease with an increase in excitation wavelength.
- The huge Stokes shift observed in DHN1 indicates the significant change in the excited state dipole moment resulting in more sensitivity towards the solvent polarity.
- The extinction coefficient has a significant role behind the origin of observed luminescence intensities apart from the minor contribution from quantum yield.
- The ProCharTS luminescence originating from the charge transfer states have lower quantum yield and the low luminescence lifetime indicating a lower efficiency of charge recombination.
- The abundant charge amino acid residue in the close proximity of the indole ring could alter the fluorescence lifetime of NATA.
- ProCharTS luminescence can alter the fluorescence emission of Trp in proteins enriched with charged amino acids, particularly Lys.

.

TH-3004_146106002

Chapter 6

Investigation of structural dynamics and functional characteristics of DHN1 and its mutants

TH-3004_146106002

6.1 Introduction:

The structural and functional relationship among proteins has been entirely associated with the structure-function paradigm, where a unique amino acid sequence encodes a unique stable 3D protein structure for a specific biological function. This structure and function paradigm has dominated in biology for more than a century regarding a functional protein or a functional protein domain. In contrast to the structure-function paradigm, Intrinsically Disordered Proteins (IDPs) gained attention due to their critical cellular role without any globular/3D structure. They established a new regime in the protein kingdom during the past decade. Despite having heterogeneous ensemble of conformations, IDPs have been shown to have significant role in key cellular processes such as cell signaling, recognition and regulation[20–22,26,69,92,212–214]. IDPs are distributed among all the organisms from simple single-cell organisms to complex organisms, plants and animals.

It has been reported that 70 % of signaling proteins are unstructured or have an intrinsically disorder region[32]. A plethora of studies reported that these IDPs have an important role in stress tolerance in the plant[215]. Recent studies show various classes of protein that are upregulated in various abiotic stressful conditions. One such protein family is Late Embryogenesis Abundant (LEA); proteins of LEA family are well-known to be upregulated in abiotic stress conditions such as water desiccation, low temperature, and high salinity[32]. Dehydrins (dehydration proteins) are a member of LEA protein family[98,216]. Dehydrins are found in different intracellular locations, mainly located in the cytoplasm and nucleus. Other locations include at mitochondria, chloroplast and plasma membrane. This diverse localization of Dehydrins reflected in the various protective functions

IDPs have been well characterized for their binding promiscuity and plasticity[38]. Heat resistance characteristics of IDPs were reported in several studies and has been successfully used in their purification[217–219]. Further, being an LEA group protein, stability and resistance to high temperature of DHN1 has various functional advantages[220–223]. We previously reported the structural changes in DHN1 at high temperature, pH and high salt conditions[199]. The structural transition in other dehydrins due to binding with the lipid vesicles and SDS micelles were reported in previous studies[94,114,221,224,225].

TH-3004_146106002

In this chapter, we studied the DHN1 from *Zea mays* protein in presence of trifluoroethanol (TFE), a well-known helix-inducing solvent which is thought to mimic the membrane environment [226,227]; at high temperature by using circular dichroism. Investigations are also performed in presence of SDS in monomer and at micellar concentration to understand the role conserved segments and N-terminal in DHN1 protein folding. DHN1 and its mutant DHN1 CW1 and DHN1 W3 in presence of SDS were studied by using the steady state fluorescence, fluorescence anisotropy, circular dichroism and time-resolved fluorescence measurement. To analyze the overall structural changes in DHN1 protein at different concentration of SDS, Förster Resonance Energy Transfer (FRET) was carried out. Further, the cryoprotection and heat protection activity of DHN1 and its mutants were measured.

TH-3004_146106002

6.2 Results and Discussion:

6.2.1 Generation of DHN1 mutants

We generated two substitution mutants DHN1 CW1 and DHN1 W3 on the basis of the position of K-segment in DHN1. Mutation scheme is given in chapter 3 **(Figure 3.1).** Single cysteine and tryptophan in DHN1 CW1 were incorporated to explore the structural dynamics of DHN1 in the presence of SDS. Single tryptophan is incorporated in DHN1 W3 at the N-terminal of DHN1 to divulge the role of the N-terminal during the structural transition of DHN1. Structural information was revealed individually by single tryptophan and external probe labeled on single cysteine (dansyl).

6.2.2 DHN1 gains ordered structure at the higher temperature

Heat stability and the essential function at elevated temperature have been reported in many IDPs[52,109]. DHN1 belongs to the LEA protein family, which accumulates during the late stages of seed development in plants[98,216]. To explore the stability and the secondary structure transitions of DHN1 at elevated temperature, far-UV CD spectra were recorded for heating DHN1 and its mutants (dissolved in water, 7 μM) from room temperature 25 °C to 90 °C and cooling back to room temperature from 90 °C. **Figure 6.1** illustrate the far-UV CD spectra of DHN1, DHN1 CW1 and DHN1 W3 upon heating and cooling back to room temperature.

DHN1, DHN1 CW1 and DHN1 W3 had a negative minimum around 198 nm and marginal negative shoulder at 222 nm, suggesting a highly unfolded structure in the native state. Subsequently, as temperature increased from 25 °C to 90 °C, disorder to order conformation transitions were observed, as shown in **Figure 6.3 A and B**. DHN1, DHN1 CW1 and DHN1 W3 shows the similar not identical behavior at higher temperature.

TH-3004_146106002

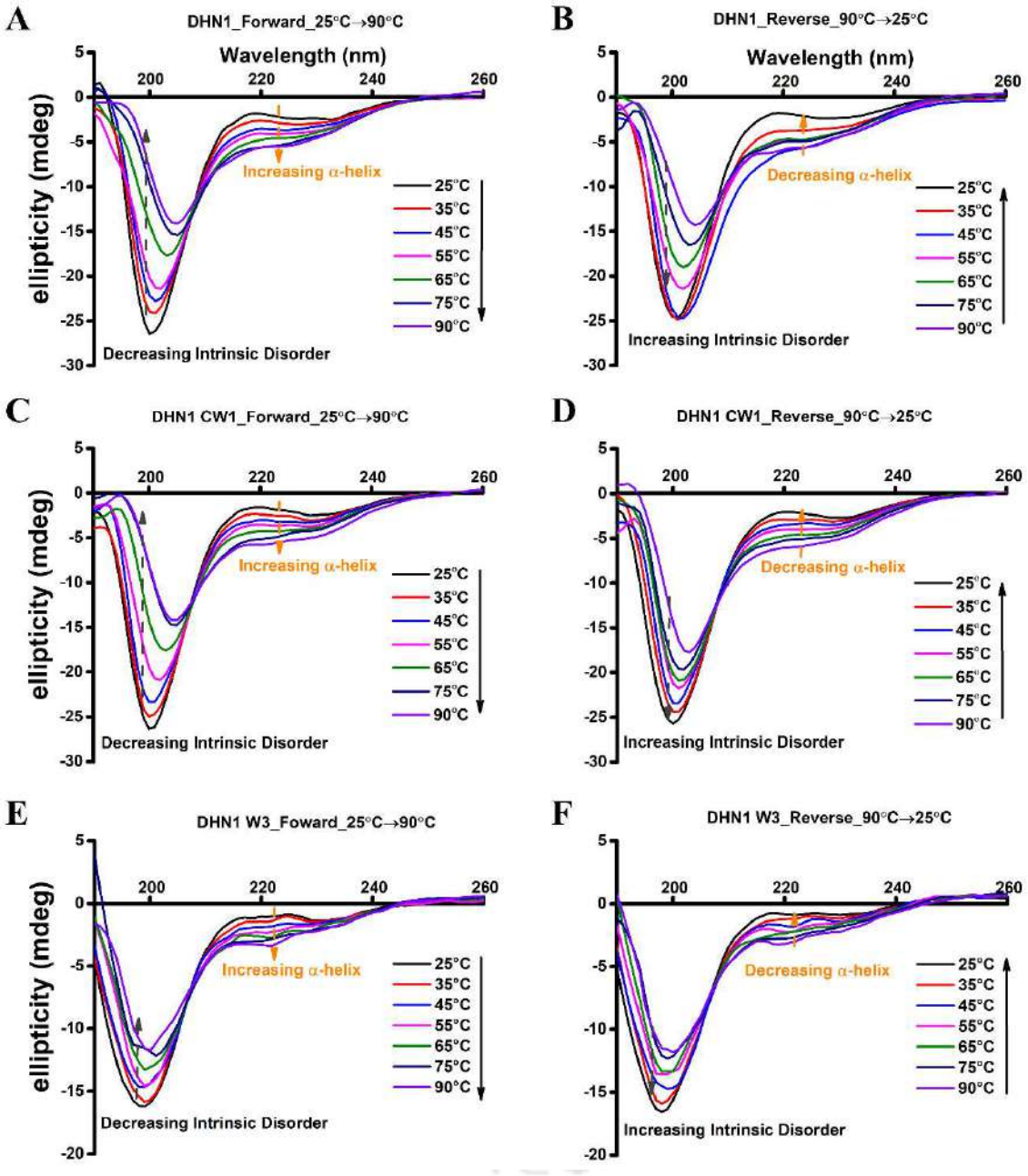


Figure 6.1: Effect of Temperature on Secondary Structure of DHN1. Far-UV CD spectra of **(A)** DHN1 Forward; **(B)** DHN1 Reverse; **(C)** DHN1 CW1 Forward; **(D)** DHN1 CW1 Reverse; **(E)** DHN1 W3 Forward; and **(F)** DHN1 W3 reverse at temperature 25 °C – 95 °C. All protein were dissolved in water at final concentration at 7 µM. Each spectrum is an average of 5 scans.

TH-3004_146106002

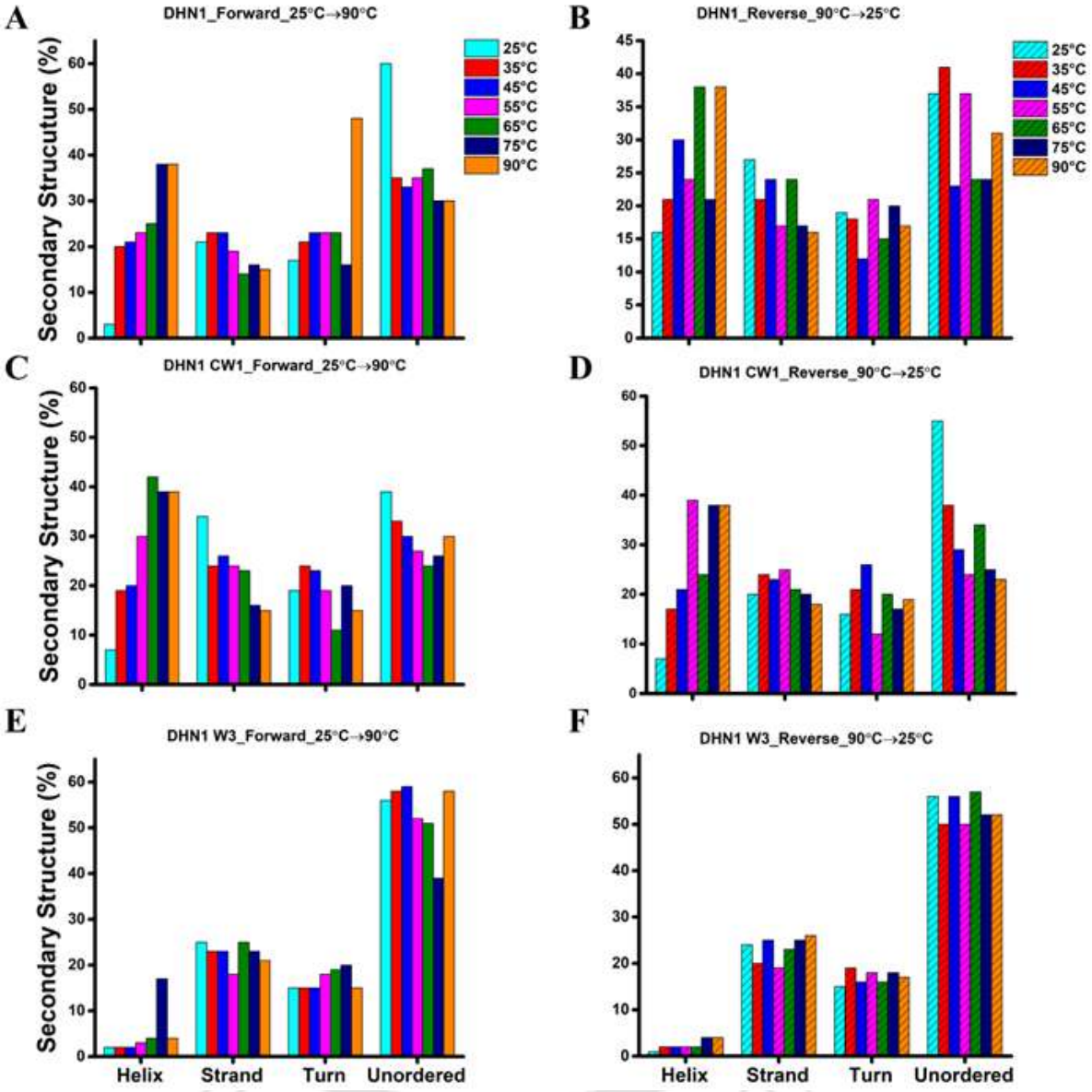


Figure 6.2: Secondary Structure content of DHN1 and Its mutants. (**A**) DHN1 Forward; (**B**) DHN1 Reverse; (**C**) DHN1 CW1 Forward; (**D**) DHN1 CW1 Reverse; (**E**) DHN1 W3 Forward; and (**F**) DHN1 W3 reverse at temperature 25 °C – 95 °C. All proteins were dissolved in water with final concentration of 7 μM.

TH-3004_146106002

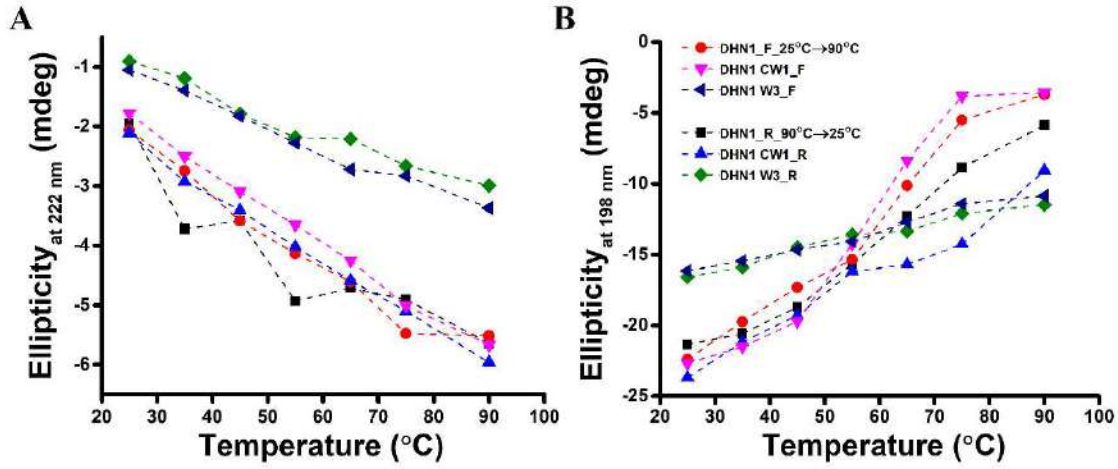


Figure 6.3: Change in ellipticity of DHN1 and its mutant. (**A**) at 222 nm; and (**B**) at 198 nm at temperature 25 °C – 95 °C (both heating up to 90°C and cooling back to 25 °C, 5-minute halt at 90 °C before cooling back to 25 °C).

To calculate the secondary structure content at different temperature, CD spectra of DHN1 and its mutant DHN1 CW1 and DHN1 W3 were analyzed using the CDSSTR program at DichroWeb. **Figure 6.2 A, C and E** displayed the change in secondary structural content upon heating from room temperature 25 °C to 90 °C of DHN1, DHN1 CW1 and DHN1 W3, respectively. Similarly, **Figure 6.2 B, D and F** represent the change in secondary structure content upon cooling back to room temperature of DHN1, DHN1 CW1 and DHN1 W3, respectively. DHN1 and its mutant DHN1 CW1 shows the substantial increase in α-helix content and decrease in disorder content upon heating. While DHN1 W3 displayed a slight change in the secondary structure content of protein. Fitted CD curves are shown in **Figure A4, A5 and A6 of appendix**. These results suggest the structural plasticity of DHN1 upon varying the temperature. Furthermore, the complete reversibility in secondary structure is also visible in DHN1 protein. These heat-induced structural changes could play an important role in stress tolerance.

6.2.3 Structural transitions in DHN1 in presence of helix inducer TFE

Structural plasticity and binding promiscuity are well characterized in IDPs[92,175,228–230]. TFE is a well-known helix inducer that has been reported in several studies related to structural transitions in IDPs[226,227]. To investigate the propensity of ordered secondary structure

TH-3004_146106002

formation of DHN1 protein, far-UV CD spectra of DHN1, DHN1 CW1 and DHN1 W3 were recorded in presence of TFE, as shown in **Figure 6.4 A, B and C,** respectively.

Figure 6.5 A, B, C and D represent the secondary structure content of DHN1, DHN1 CW1 and DHN1 W3, respectively, with increasing concentration of TFE obtained from CDSSTR program at DichroWeb. Fitted CD curves are shown in the **Figure A7, A8 and A9 of appendix**.

Secondary structural transitions from disordered to ordered fraction as increase in TFE concentration are clearly visible in CD spectra of DHN1 and its mutant DHN1 CW1 and DHN1 W3, **Figure 6.4 A, B and C**. The α-helical conformation in DHN1 and its mutant DHN1 CW1 and DHN1 W3 at physiological pH 7.4 (dissolved in water) have ≤5% of helix, which increased tremendously ≥80% upon increasing TFE (10%−90%). Furthermore, **Figure 6.5 E and F** display the change in ellipticity at 222 nm (rise in α-helical conformation) and 198 nm (decrease in disordered conformation), respectively. DHN1 and its mutant DHN1 CW1 and DHN1 W3 show similar but not identical structural transitions in presence of TFE. These folding propensity of DHN1 clearly demonstrate towards its tendency of α-helix formation upon interaction with binding partner and plays a key role in various cellular process[220,230–234].

TH-3004_146106002

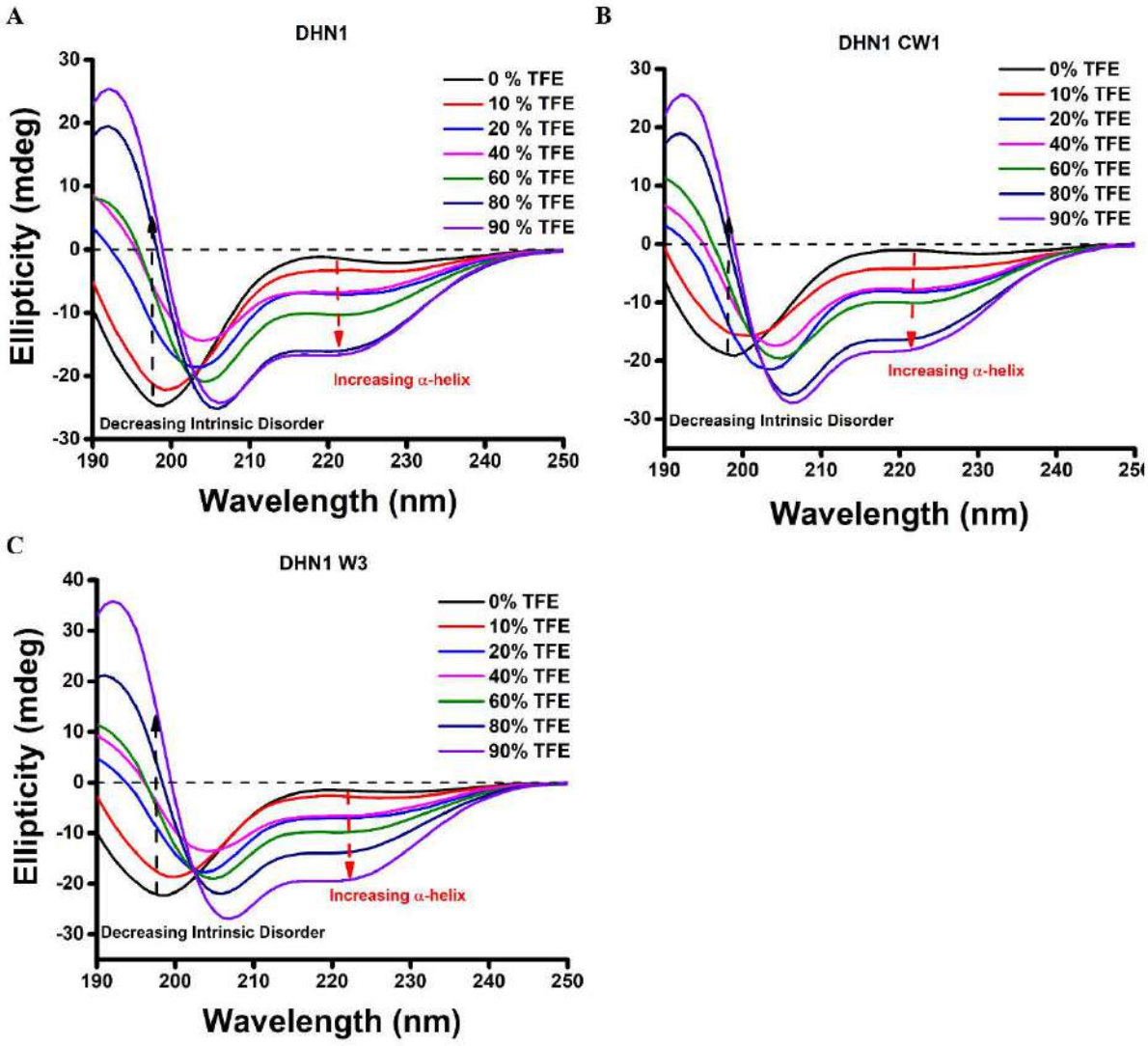


Figure 6.4: Effect of TFE on Secondary Structure of DHN1. Far-UV spectra of 7 µM of (**A**) DHN1; (**B**) DHN1 CW1; and (**C**) DHN1 W3 at various concentration of TFE. Each spectrum is an average of 5 scans.

TH-3004_146106002

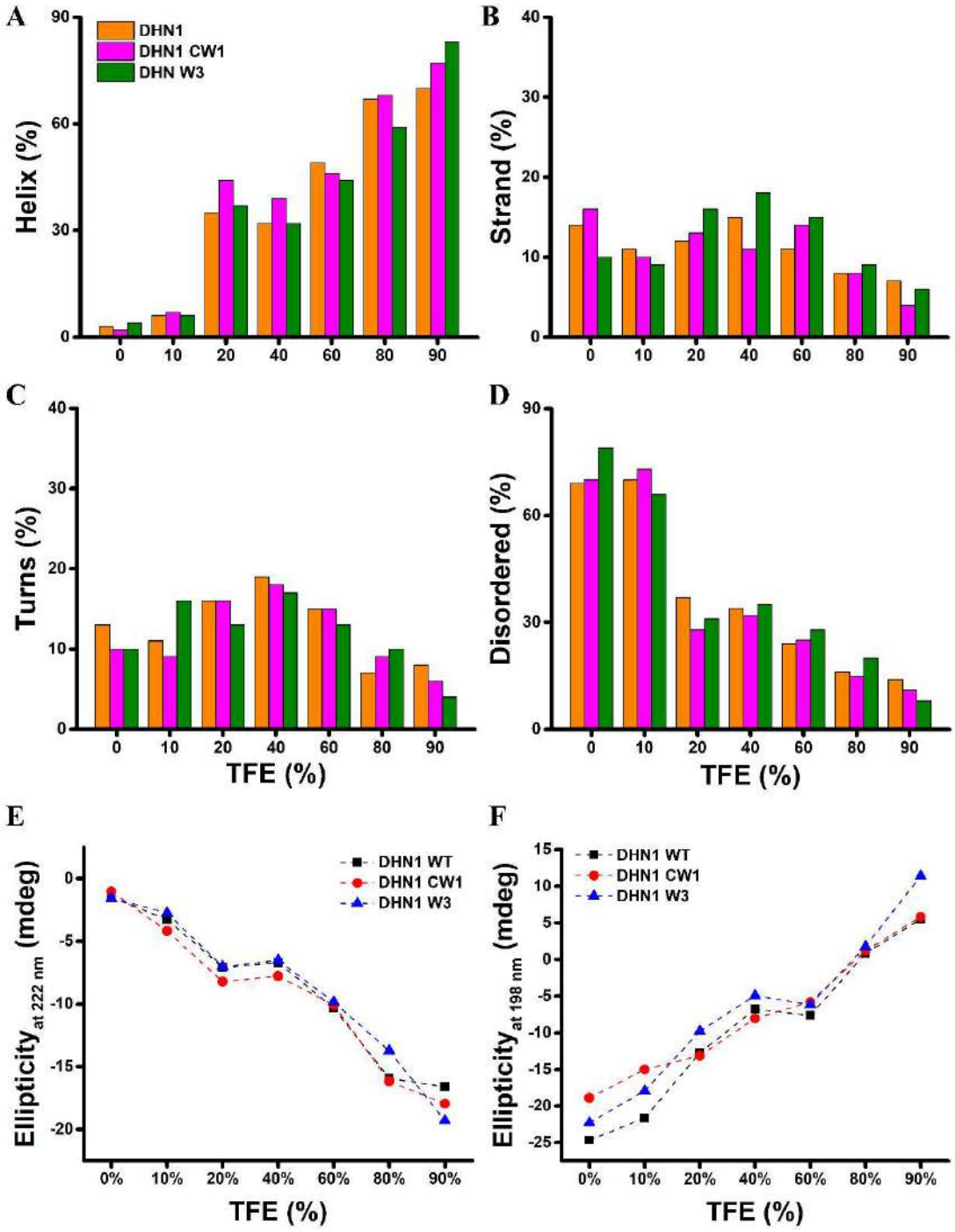


Figure 6.5: Effect of TFE on Secondary Structure of DHN1. Structure content of 7 µM of **(A)** Helix Content; **(B)** Strand Content; **(C)** Turns Content; **(D)** Disorder content; **(E)** Change in ellipticity at 222 nm; and **(F)** at 198 nm in presence of various concentration of TFE.

TH-3004_146106002

6.2.4 Effect of SDS on DHN1 and its mutant

To understand the correlation of its structural gain upon interaction with its different partner and its functional importance, we investigated the structural transitions of DHN1 protein in the presence of lipid membrane mimicking agent, SDS as monomer, and micelles. Protein folding upon interaction with a partner has been reported in several works[235–239]. For analyzing the structural transition upon binding with the SDS, far-UV CD spectra of DHN1 and its mutant in presence of SDS were recorded, as displayed in **Figure 6.6 A, B and C**.

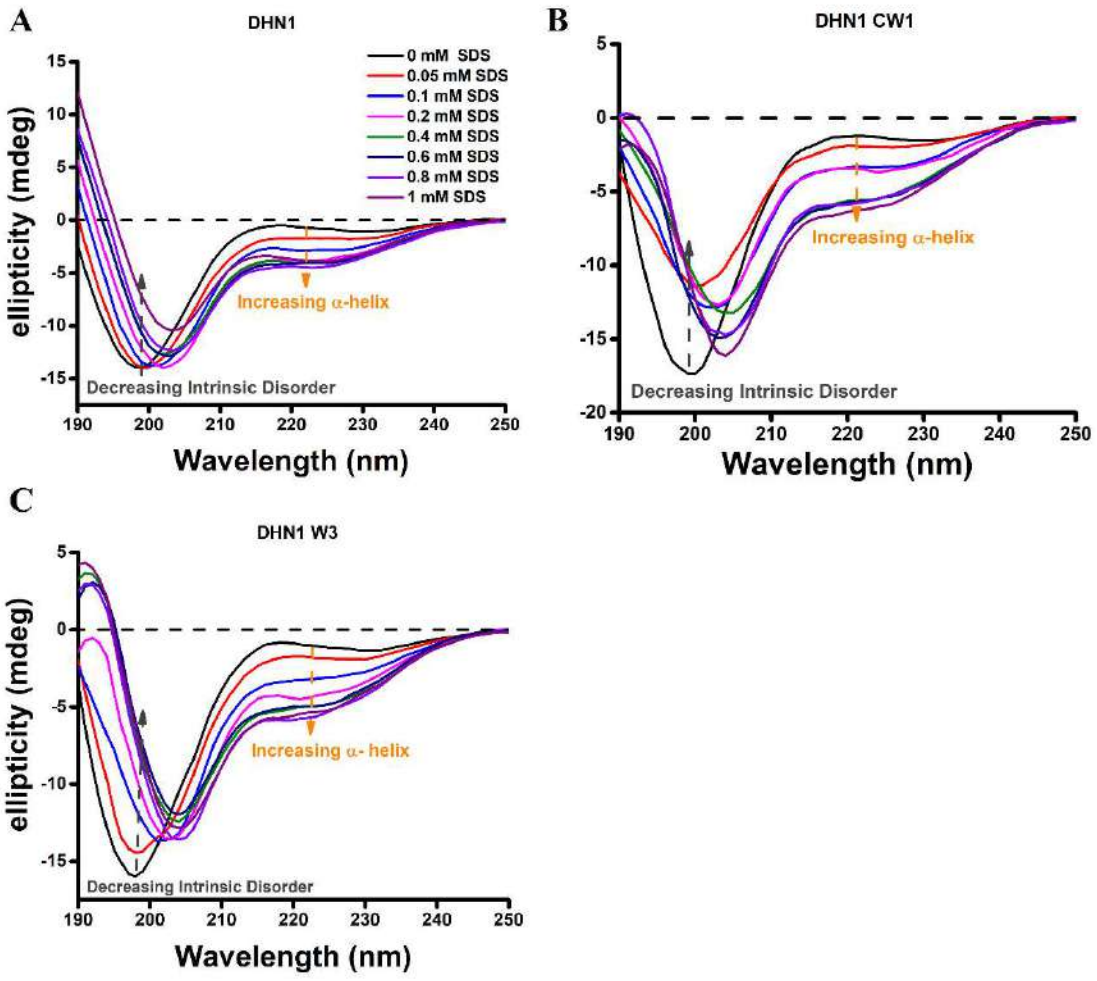


Figure 6.6: Effect of SDS on Secondary Structure of DHN1. Secondary Structure content of 7 µM of **(A)** DHN1; **(B)** DHN1 CW1; and **(C)** DHN1 W3 at various concentration of SDS (0.05 mM −1 mM). Each spectrum is an average of 5 scans.

TH-3004_146106002

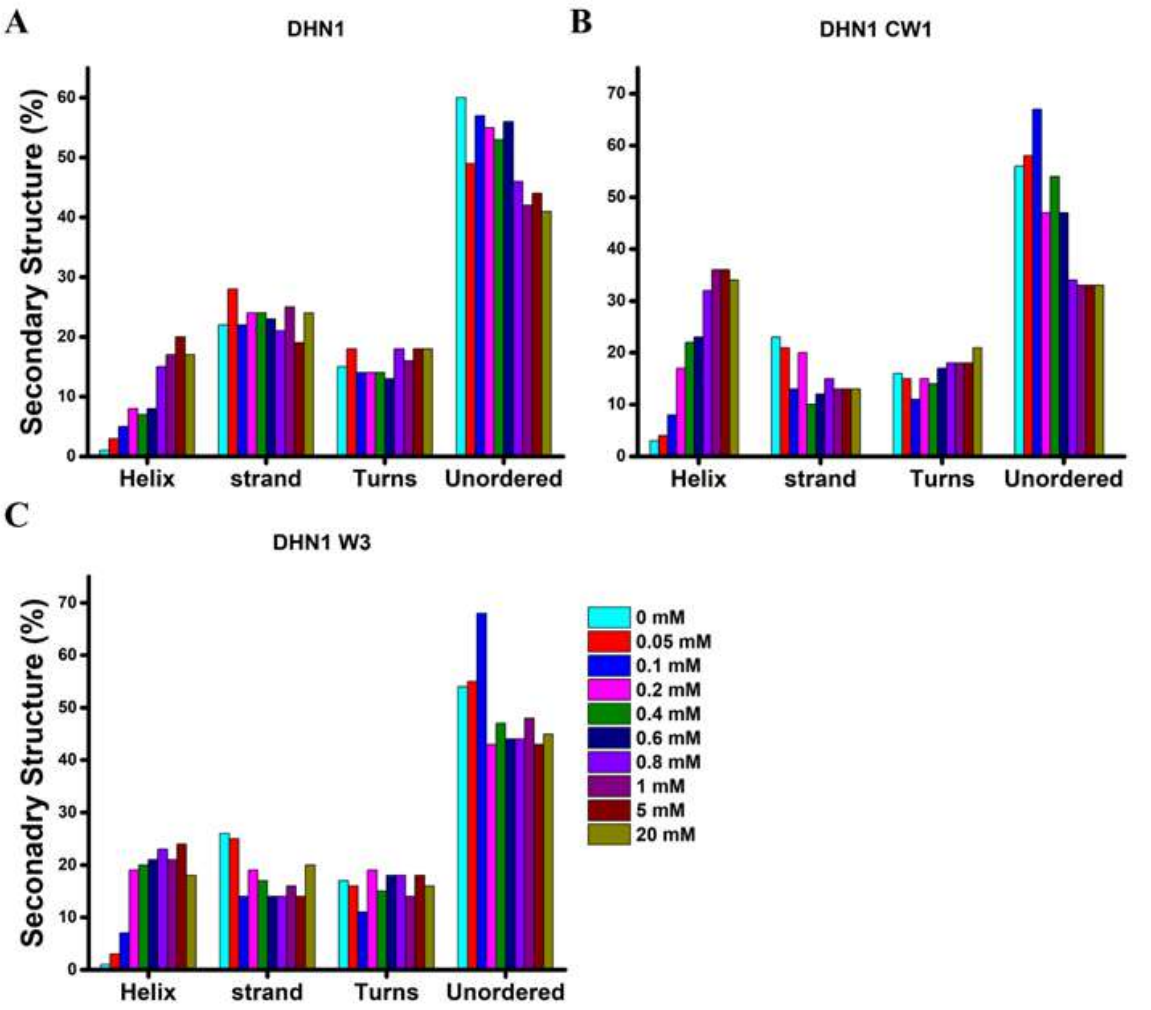


Figure 6.7: Effect of SDS on Secondary Structure of DHN1. Secondary Structure content of 7 µM of (**A**) DHN1; (**B**) DHN1 CW1; and (**C**) DHN1 W3 at various concentration of SDS (0.05 mM−20 mM).

Figure 6.7 A, B and C illustrate the secondary structure content of DHN1, DHN1 CW1 and DHN1 W3, respectively, at different concentration of SDS; analyzed using the CDSSTR program. Fitted CD curves are displayed in **Figure A10, A11 and A12 of appendix**. The extended IDP (native coils or native pre-molten globules) DHN1, DHN1 CW1 and DHN1 W3 secondary structure transitions were induced by the SDS (anionic surfactant). DHN1 and its mutant shows the similar but not identical structural transition in presence of SDS.

Transition from disorder to order conformation is clearly demonstrated by the change in ellipticity at 222 nm and 198 nm, as shown in **Figure 6.8 A, B and C**. SDS in monomeric form was able to induce the structural transitions in DHN1. Upon increasing concentration

TH-3004_146106002

of SDS (up to 1 mM) α-helical conformation of DHN1 increased up to 20% while disordered conformation reduced to 40% in DHN1 and its mutants. At micellar concentration of SDS like 20 mM, no significant change is observed compared to 1 mM. Our result suggests that the even monomer form of SDS (0.05 mM) induced structure in DHN1. **Table 6.1** shows the fitted parameters obtained for the change in

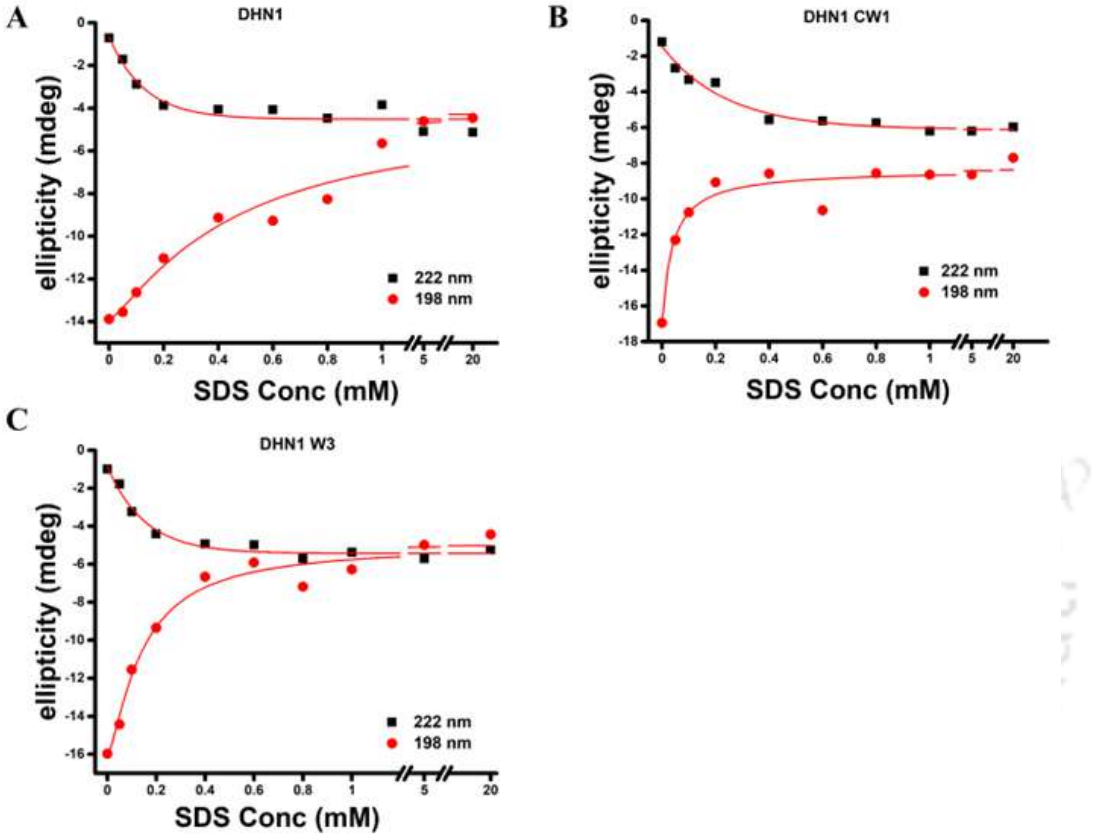


Figure 6.8: Effect of SDS on Secondary Structure of DHN1. Change in ellipticity of 7 μM (**A**) DHN1; (**B**) DHN1 CW1; and (**C**) DHN1 W3 at 222 nm (black filled square) and 198 nm (red-filled circle) at various concentration of SDS. Break in X-axis is between 1.5 to 4.5 and 5.5 to 19.5. Connected red and black lines represent the non-linear curve fitting.

Table 6.1: Fitted parameters obtained form the non-linear curve fitting of the change in the ellipticity at 222 nm and 198 nm as function of SDS.

Protein	Exponential decay fit* for change in ellipticity at 222 nm			Logistic fit# for change in ellipticity at 198 nm		
	t1	A1	R^2	x0	p	R^2
DHN1	0.13 ± 0.04	3.84 ± 0.45	0.91	0.46 ± 0.10	1.22 ± 0.32	0.97
DHN1 CW1	0.24 ± 0.04	4.63 ± 0.31	0.93	0.04 ± 0.01	1.00 ± 0.53	0.93
DHN1 W3	0.15 ± 0.02	4.56 ± 0.28	0.98	0.14 ± 0.02	1.34 ± 0.25	0.98

* Please refer to appendix for equations

TH-3004_146106002

ellipticity at 222 nm and 198 nm as function of SDS concentration. DHN1 CW1 and DHN1 W3 were found to be more sensitive to SDS in comparison to DHN1.

6.2.4.1 Steady state fluorescence emission and anisotropy of DHN1

Steady-state fluorescence emission and anisotropy (r_{ss}) of DHN1 CW1 and DHN1 W3 tryptophan was recorded to extract the information around the indole ring and its rotational dynamics in various SDS concentration. Tryptophan emission is highly sensitive to the surrounding environmental conditions; its emission maximum in folded protein (buried inside a protein) is close to 320 nm, while in unstructured (solvent exposed) it has an emission maximum close to 350 nm [156,240]. To understand the effect of SDS on the structure and dynamics of DHN1, fluorescence emission spectra of DHN1 CW1 (10 µM), DHN1 W3, and NATA (10 µM) at different SDS concentration were

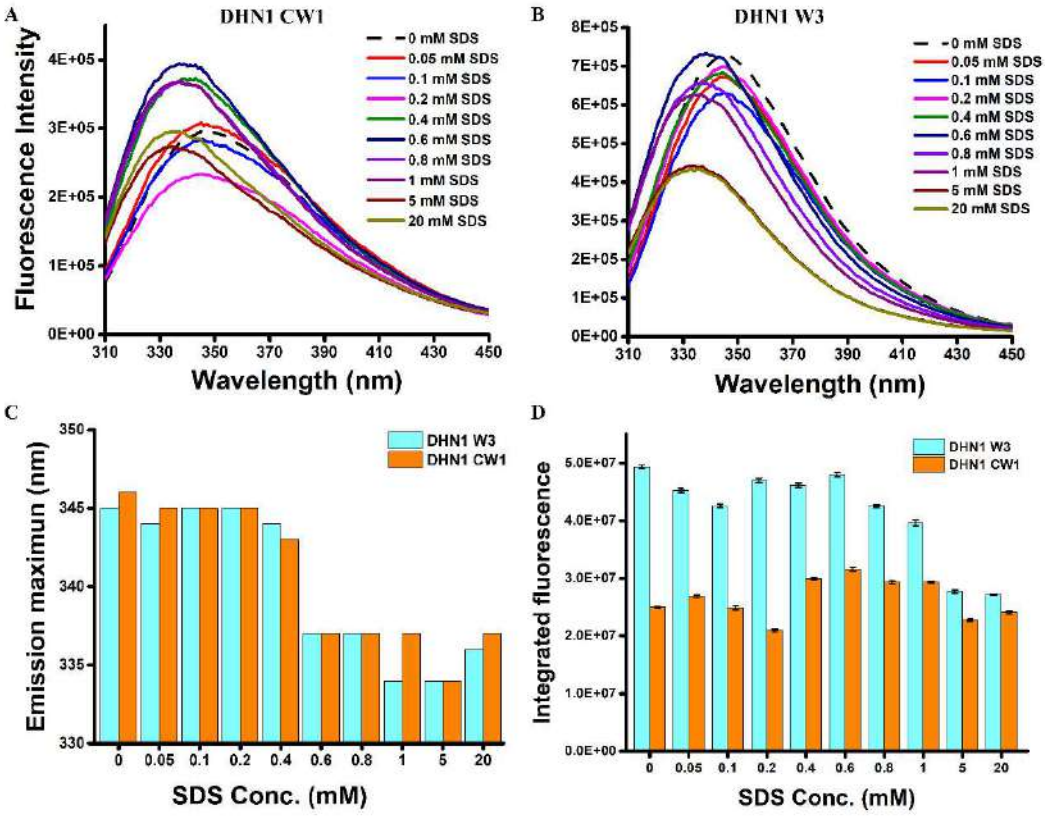


Figure 6.9: Effect of SDS on tryptophan steady-state fluorescence emission of DHN1. Fluorescence emission spectra of 10 µM of (**A**) DHN1 CW1; (**B**) DHN1 W3; and (**C**) Emission maximum of DHN1 CW1 and DHN1 W3; (**D**) Integrated fluorescence of DHN1 CW1 and DHN1 W3.

TH-3004_146106002

collected by exciting the sample at 295 nm. **Figure 6.9 A and B** shows the steady state fluorescence emission of DHN1 CW1 and DHN1 W3 tryptophan, respectively. **Figure 6.9 C and D** display the emission maxima and integrated fluorescence emission of DHN1 CW1 and DHN1 W3 tryptophan, respectively. Both DHN1 CW1 and DHN1 W3 tryptophan emission maximum was observed around 346 nm (pH 7.4, 50 mM Sodium Phosphate Buffer), while on increasing SDS concentration, the emission maximum shift to ~333 nm. The blue shift in its emission maximum indicates that the tryptophan of DHN1 CW1 and DHN1 W3 is gradually buried inside the protein core as the SDS concentration increases. Interestingly, the N-terminal of DHN1 protein also become part of the compact core of protein upon folding in the presence SDS, which is clearly indicated by the blue shift in emission maxima of DHN1 W3 tryptophan.

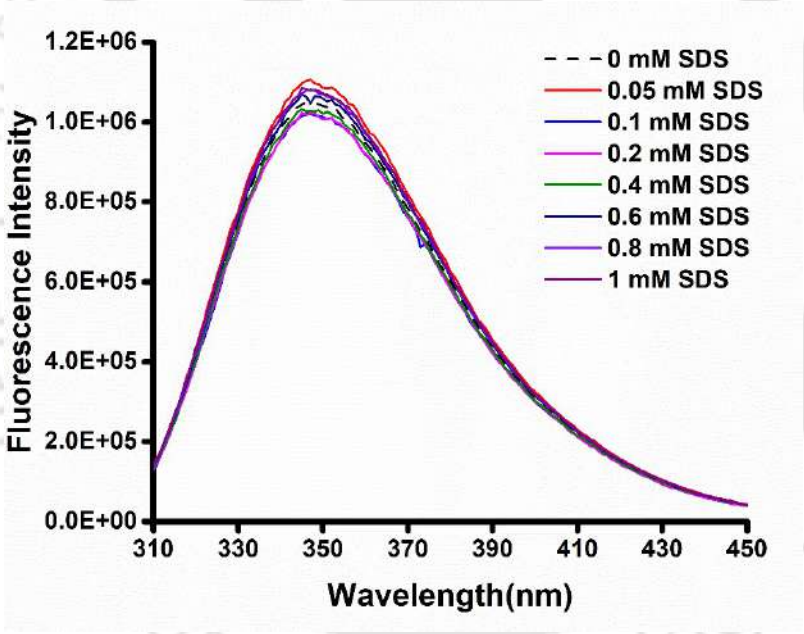


Figure 6.10: Fluorescence emission spectra of NATA in presence of SDS: Fluorescence intensity of NATA (20 µM) with different concentration of SDS upon excitation at 295 nm and emission collected from 310 – 580 nm. Slit width kept for excitation and emission is 2 nm and 10 nm, respectively.

.

Figure 6.10 displays the fluorescence emission spectra of NATA display insignificant change in its emission maximum at different concentration of SDS as it is fully exposed to the solvent. We monitored the fluorescence anisotropy of DHN1 CW1 and DHN1 W3 tryptophan in presence of SDS. **Figure 6.11A** represents the DHN1 CW1 and DHN1 W3

TH-3004_146106002

tryptophan fluorescence anisotropy with increasing SDS concentration (0.05 mM−20 mM). Both DHN1 CW1 and DHN1 W3 shows similar but not identical behavior. DHN1 CW1 mutant reveals a structural transition at lower concentration of SDS compared to DHN1 W3 mutant. So, N-terminal region folds later in DHN1 in presence of SDS, compared to C-terminal. Change in the anisotropy of NATA was insignificant in different concentration of SDS (shown in appendix, **Figure A13**). Anisotropy increased with increase in SDS concentration hints that the rotation of the tryptophan slows down due to its incorporation into the hydrophobic core. Similarly, **Figure 6.11B** shows the fluorescence anisotropy of dansyl labeled on cysteine[62] of DHN1 CW1 in presence of SDS. An increase in dansyl anisotropy is associated with the slow rotation of dansyl due to the folding of DHN1 protein.

Tryptophan and dansyl probe anisotropy in DHN1 CW1 and DHN1 W3 suggests that the tryptophan and dansyl rotation hindered due to local folding of the DHN1. This result also points toward the crucial role of the N-terminal in DHN1 folding.

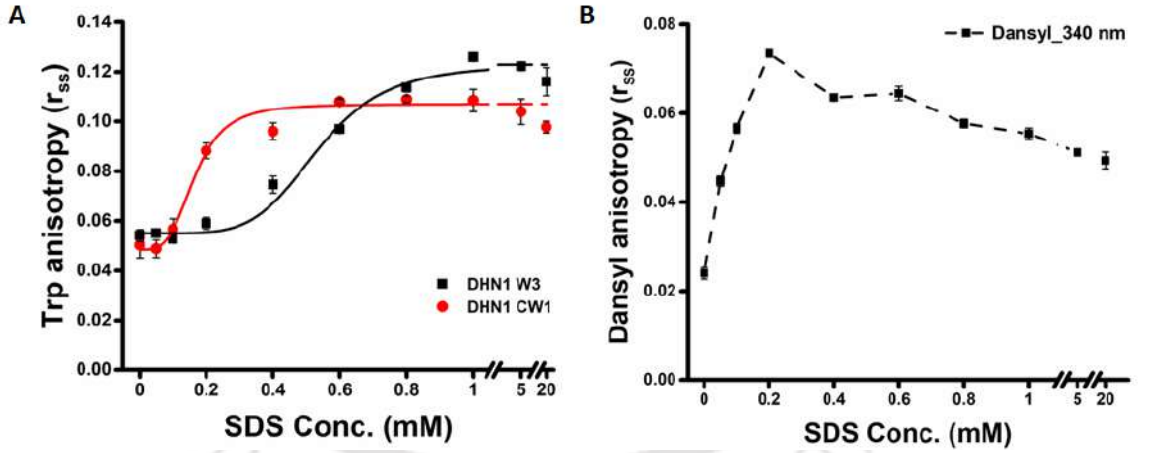


Figure 6.11: Effect of SDS on Steady state Fluorescence anisotropy of DHN1. (A) Steady state Trp fluorescence anisotropy of DHN1 CW1 and DHN1 W3; and (**B**) steady state fluorescence anisotropy of dansyl (IAEDANS) labeled on Cysteine [122] of DHN1 CW1 at various concentration of SDS. Standard deviations, n=3 at each point are shown. Connected red and black lines represent the non-linear curve fitting (refer **appendix Table A1** for parameters).

6.2.4.2 Fluorescence lifetime and FRET measurement in presence of SDS

The structural transition between the large ensemble of disordered conformation is critical to understanding the dynamics of DHN1 protein. Several approaches to utilize FRET to

TH-3004_146106002

understand the dynamics of proteins have been developed over the past few years [45,241–243]. Structural information of an IDP is challenging to obtain under native conditions. To extract the precise structural information, we monitored the Intramolecular Förster Resonance Energy (FRET) distance between the Cys^{62} (dansyl) and Trp^{122} (indole) of DHN1 CW1 at different SDS concentration. DHN1 CW1 Cysteine was covalently labeled with external

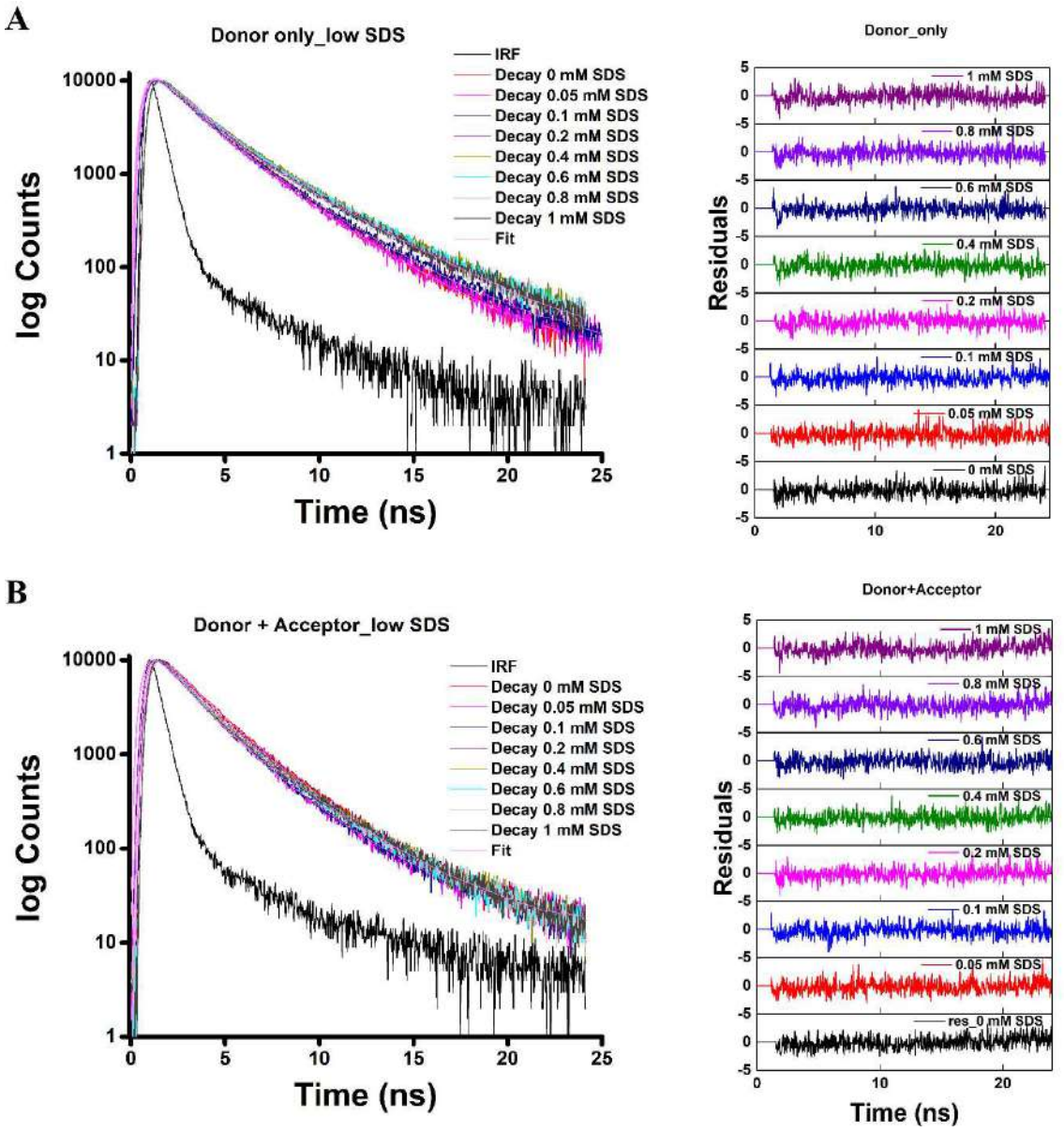


Figure 6.12: Effect of monomer SDS on the distance between Cys^{62} (dansyl)- Trp^{122} (Indole) in DHN1 CW1. Fitted time-resolved fluorescence intensity decay of 10 μM DHN1 CW1 tryptophan (donor) excited at λ_{ex} 295 nm in presence (**A**) donor only (**B**) donor+ acceptor, at different concentration of SDS (0−1 mM). Residuals for the fit is shown in right panel.

TH-3004_146106002

probe dansyl (IAEDANS) with the dye/protein ratio of ~0.9. Time-resolved fluorescence intensity decay of DHN1 CW1 tryptophan was measured in the presence and absence of acceptor (dansyl) at different SDS concentration (0.05 mM−1 mM). **Figure 6.12 A and B** display the fluorescence intensity decay profile of DHN1 CW1 tryptophan in the absence and presence of acceptor, respectively, in the presence of SDS from 0.05 mM to 1 mM concentration**.** Similarly, the fluorescence intensity decay profile of DHN1 CW1 tryptophan measured in the absence and presence of acceptor at SDS concentration from 1 mM to 20 mM is shown in **Figure A14 of the appendix**.

In the presence of SDS, DHN1 CW1 tryptophan shows the faster decay in fluorescence intensity. This indicates significant energy transfe from tryptophan to the acceptor. Further, FRET efficiency and distance between cysteine^{62} and tryptophan^{122} of DHN1 CW1 were calculated using the equation 2.48 and 2.46, respectively.

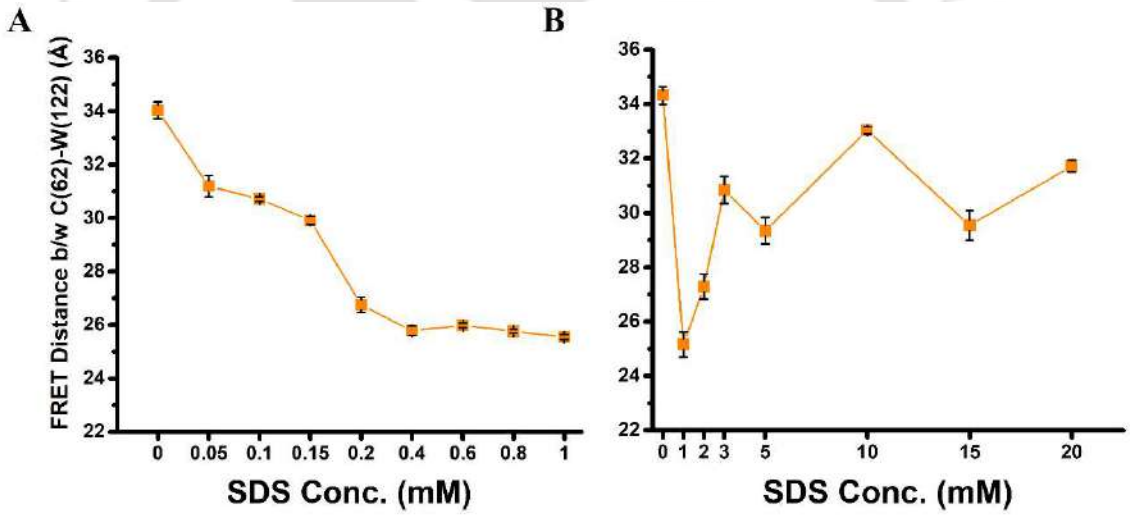


Figure 6.13. Effect of SDS on the distance between Cys^{62} (dansyl)- Trp^{122} (Indole) in DHN1 CW1. FRET distances between tryptophan^{122} (donor) and dansyl labeled cysteine^{62} (acceptor) of DHN1 CW1 **(A)** at lower SDS concentration (0−1 mM) range (**B)** at high SDS concentration (0−20 mM) range. Standard deviations at each point are shown.

Figure 6.13 A and B show the FRET distance between tryptophan^{122} (donor) and dansyl labelled cysteine^{62} (acceptor) of DHN1 CW1 at different concentration of SDS (0.05 mM−1 mM). **Table 6.2 A and B** shows the calculated lifetime of tryptophan, FRET efficiency and distance in the presence of SDS 0.05−1 mM and 1−20 mM, respectively. The Förster

TH-3004_146106002

distance for the TRP-DANS FRET pair is 22 Å [156]. Distance calculated in native DHN1 CW1 was 34 Å. In the presence of SDS, the FRET distance between tryptophan[122] and dansyl labelled cysteine[62] was decreased to 25 Å. This decrease hints toward the change in native conformation of DHN1 CW1, which is also corroborated by the change in secondary structure in far-UV CD Spectra. At a very low concentration of SDS, Tryptophan[122] which is present in between two K-segment and the Cysteine[62] which is near the S-segment are coming to close proximity which suggests the folding of K- segment. Subsequently, upon increasing the concentration of SDS above 1 mM, there was slight increase in FRET distance were observed as shown in **Figure 6.13 B.**

TH-3004_146106002

Table 6.2 A: Average fluorescence lifetime of DHN1 CW1 tryptophan (in presence and absence of acceptor), FRET efficiency and distance between Cys^{62}-Trp^{122} at different concentration of SDS (monomer) (0.05−1 mM). The standard deviation for n = 3 is shown in brackets.

SDS Conc. (mM)		χ^2 reduced	τ_1 (ns)	α_1	τ_2 (ns)	α_2	τ_m (ns)	Average FRET Efficiency	Average Distance b/w Cys^{62}-Trp^{122} (Å)
0	Donor	1.00	2.97 (0.055)	0.413	1.45 (0.41)	0.587	2.1 (0.037)	0.068 (0.002)	34.03 (0.32)
	Donor+ Acceptor	1.12	3.02 (0.479)	0.306	1.54 (0.161)	0.694	1.96 (0.048)		
0.05	Donor	1.14	2.87 (0.111)	0.38	1.40 (0.097)	0.615	2.02 (0.061)	0.110 (0.003)	31.19 (0.40)
	Donor+ Acceptor	1.15	2.4 (0.108)	0.423	1.25 (0.092)	0.577	1.80 (0.022)		
0.1	Donor	1.10	2.94 (0.045)	0.45	1.34 (0.06)	0.541	2.08 (0.042)	0.119 (0.005)	30.71 (0.09)
	Donor+ Acceptor	1.03	2.74 (0.263)	0.222	1.47 (0.06)	0.778	1.83 (0.009)		
0.15	Donor	1.10	3.11 (0.038)	0.45	1.29 (0.053)	0.547	2.09 (0.033)	0.137 (0.010)	29.91 (0.16)
	Donor+ Acceptor	1.03	2.65 (0.02)	0.316	1.44 (0.014)	0.684	1.81 (0.006)		
0.2	Donor	1.05	3.74 (0.016)	0.304	1.67 (0.022)	0.696	2.33 (0.002)	0.236 (0.018)	26.75 (0.28)
	Donor+ Acceptor	1.10	2.53 (0.055)	0.478	1.01 (0.066)	0.522	1.78 (0.029)		
0.4	Donor	1.01	3.8 (0.094)	0.312	1.62 (0.065)	0.688	2.36 (0.025)	0.278 (0.011)	25.79 (0.18)
	Donor+ Acceptor	1.03	2.77 (0.065)	0.379	1.10 (0)	0.621	1.70 (0.012)		
0.6	Donor	1.00	3.73 (0.004)	0.324	1.55 (0.007)	0.676	2.29 (0.025)	0.269 (0.013)	25.98 (0.09)
	Donor+ Acceptor	1.07	2.82 (0.039)	0.308	1.13 (0.06)	0.692	1.67 (0.015)		
0.8	Donor	1.09	3.63 (0.205)	0.320	1.45 (0.029)	0.680	2.17 (0.005)	0.278 (0.008)	25.79 (0.06)
	Donor+ Acceptor	1.09	2.83 (0.101)	0.295	1.00 (0.073)	0.705	1.59 (0.043)		
1	Donor	1.10	3.69 (0.196)	0.299	1.436 (0.132)	0.701	2.22 (0.049)	0.290 (0.017)	25.56 (0.09)
	Donor+ Acceptor	1.01	2.74 (0.066)	0.331	0.99 (0.058)	0.669	1.58 (0.031)		

TH-3004_146106002

Table 6.2 B: Average fluorescence lifetime of DHN1 CW1 tryptophan (in presence and absence of acceptor), FRET efficiency and distance between Cys^{62}-Trp^{122} at different concentration of SDS (1−20 mM). The standard deviation for n = 3 is shown in brackets.

SDS Conc. (mM)		χ_R^2	τ_1 (ns)	α_1	τ_2 (ns)	α_2	τ_m (ns)	Average FRET Efficiency	Average Distance b/w Cys^{62}-Trp^{122} (Å)
0	Donor	1.00	1.19 (0.010)	0.480	2.93 (0.027)	0.520	2.11 (0.01)	0.065 (0.002)	34.32 (0.57)
	Donor+ Acceptor	1.08	1.14 (0.211)	0.469	2.70 (0.228)	0.986	1.99 (0)		
1	Donor	1.05	1.35 (0.020)	0.651	3.75 (0.037)	0.349	2.30 (0.03)	0.309 (0.017)	25.16 (0.71)
	Donor+ Acceptor	1.12	1.11 (0.004)	0.682	3.16 (0.004)	0.318	1.59 (0.10)		
2	Donor	1.04	1.18 (0.023)	0.662	3.48 (0.011)	0.338	2.04 (0.10)	0.216 (0.012)	27.28 (0.71)
	Donor+ Acceptor	1.16	1.12 (0.029)	0.688	3.14 (0.001)	0.312	1.6 (0.13)		
3	Donor	1.09	1.17 (0.005)	0.652	3.42 (0.035)	0.348	2.037 (0.06)	0.116 (0.030)	30.84 (0.75)
	Donor+ Acceptor	1.19	1.18 (0.008)	0.711	3.22 (0.011)	0.289	1.8 (0.17)		
5	Donor	1.06	1.22 (0.009)	0.673	3.43 (0.039)	0.327	2.061 (0.05)	0.151 (0.039)	29.34 (0.74)
	Donor+ Acceptor	1.08	1.10 (0.011)	0.704	3.01 (0.059)	0.296	1.75 (0.20)		
10	Donor	1.09	1.23 (0.001)	0.672	3.48 (0.047)	0.328	2.094 (0.07)	0.080 (0.001)	33.04 (0.29)
	Donor+ Acceptor	1.27	1.21 (0.076)	0.717	3.21 (0.151)	0.283	1.9 (0.05)		
15	Donor	1.20	1.23 (0.002)	0.695	3.5 (0.046)	0.305	1.99 (0.24)	0.146 (0.057)	29.54 (0.79)
	Donor+ Acceptor	1.22	1.17 (0.008)	0.751	3.21 (0.069)	0.249	1.78 (0.05)		
20	Donor	1.15	1.24 (0.015)	0.686	3.45 (0.058)	0.314	1.96 (0.005)	0.101 (0.010)	31.73 (0.23)
	Donor+ Acceptor	1.15	1.19 (0.074)	0.685	3.10 (0.193)	0.261	1.75 (0.005)		

TH-3004_146106002

To further gain information of the N-terminal of DHN1 protein, the time-resolved fluorescence lifetime of its mutant DHN1 W3 tryptophan was measured in the presence of different concentration of SDS (0.05 mM −1 mM). **Figure 6. 14** displays the time-resolved fluorescence intensity decay of 10 µM DHN1 W3 tryptophan in presence of SDS.

Analysis of fluorescence intensity decay reveal presence of two lifetimes. The values obtained from the fitted fluorescence intensity decay curve are shown in **Figure 6.15 A, B and C**. The τ_1 with α_1 ~0.35 remains unchanged while the decrease is noticeable in τ_2 with α_2 values ~0.65. The mean lifetime of DHN1 W3 tryptophan decreased as increase in SDS concentration. **Table 6.3** display the fluorescence lifetime values of DHN1 W3 tryptophan in presence of SDS (0.05 mM -1 mM).

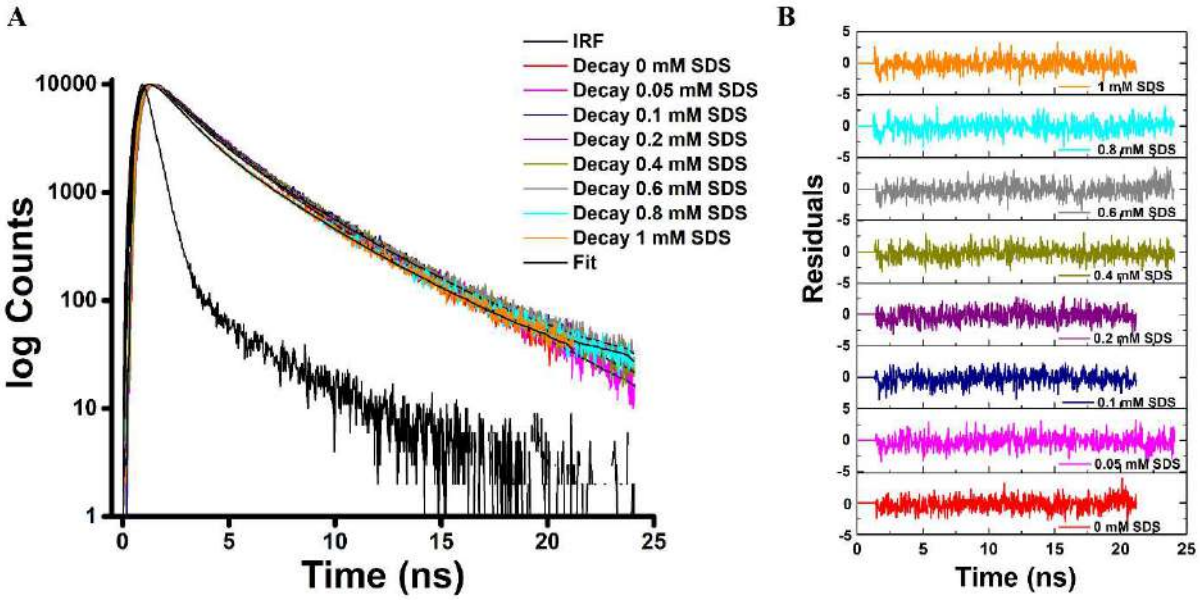


Figure 6.14: (A) Fitted time-resolved fluorescence intensity decay profile of 10 µM DHN1 W3 tryptophan (λ_{ex} of 295 nm) in presence of different concentration of SDS (0−1 mM). (**B**) Residuals for the fit.

.

TH-3004_146106002

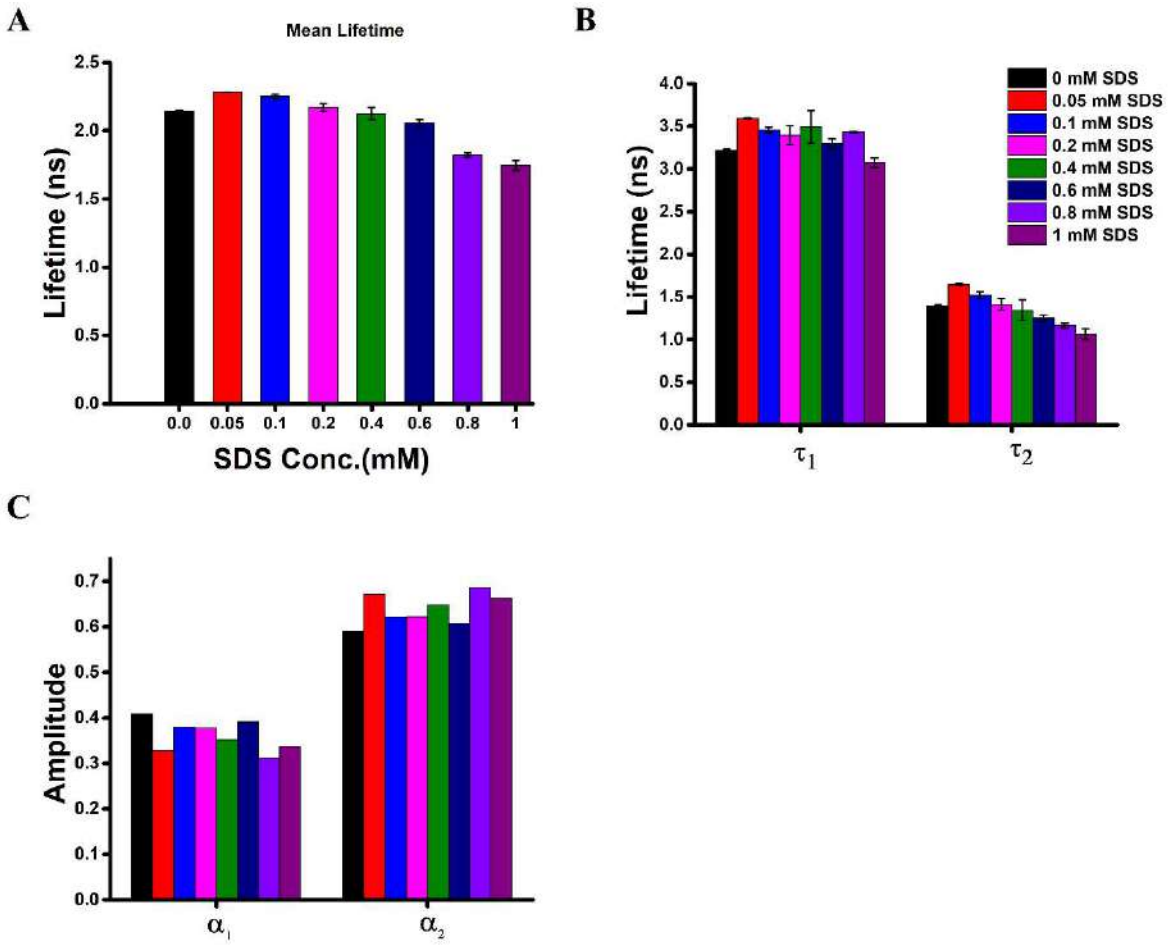


Figure 6.15: Effect of SDS on DHN1 W3: Time-resolved fluorescence at λ_{ex} of 295 nm of 10 µM DHN1 W3 (**A**) Mean lifetime and (**B**) different τ values; and (**C**) different α values at different concentration of SDS. The data represents the average of 3 individual measurement with error bar corresponds to standard deviation.

TH-3004_146106002

Table 6.3 Parameters obtained from the two-exponential fit for the fluorescence intensity decay of DHN1 W3 tryptophan in presence of different concentration of SDS (0.05 mM −1 mM). Values in bracket (for τ_i and τ_m) represents the standard deviation observed in two or more independent experiments, while α_i values are from the best fit.

SDS Conc. (mM)	α_1	α_2	τ_1 (ns)	τ_2 (ns)	τ_m (ns)	χ^2_R
0	0.42	0.58	3.27 (0.013)	1.39 (0.019)	2.14 (0.003)	1.04
0.05	0.32	0.68	3.60 (0.009)	1.65 (0.008)	2.28 (0)	1.08
0.1	0.36	0.63	3.46 (0.030)	1.52 (0.039)	2.25 (0.016)	1.03
0.2	0.42	0.58	3.40 (0.109)	1.41 (0.066)	2.17 (0.029)	1.00
0.4	0.42	0.58	3.49 (0.187)	1.35 (0.118)	2.13 (0.044)	1.10
0.6	0.38	0.62	3.30 (0.054)	1.25 (0.038)	2.06 (0.025)	1.12
0.8	0.31	0.69	3.2 (0.231)	1.17 (0.029)	1.82 (0.016)	1.06
1	0.35	0.65	3.07 (0.056)	1.07 (0.063)	1.75 (0.037)	1.01

To gain further insights into the effect of SDS on DHN1 protein, fluorescence lifetime measurement of dansyl labeled DHN1 CW1 was performed at different SDS concentration (1 mM -20 mM). **Figure 6.16** shows the intensity decay profile of dansyl labeled DHN1 CW1 at different SDS concentration. Modest change was observed in fluorescence intensity decay of dansyl labeled DHN1 CW1. Analysis of fluorescence intensity decay display the presence of two lifetimes. Mean fluorescence lifetime (τ_m) of dansyl labeled DHN1 CW1 decreased from 13.35 to 12.18 ns with increase in SDS concentration. Interestingly, the mean lifetime remains almost constant after 3 mM of SDS concentration as shown in **Figure**

TH-3004_146106002

6.17. **Table 6.4** shows the lifetime values of dansyl labeled DHN1 CW1 at different concentration of SDS, obtained after fitting the fluorescence intensity decay profile.

The overall results suggests that the DHN1 and its mutant DHN1 CW1 and DHN1 W3 are sensitive to very low concentration of SDS, further increase in concentration above the micellar concentration (8.3 mM) of SDS does not cause much influence on the DHN1 secondary structure.

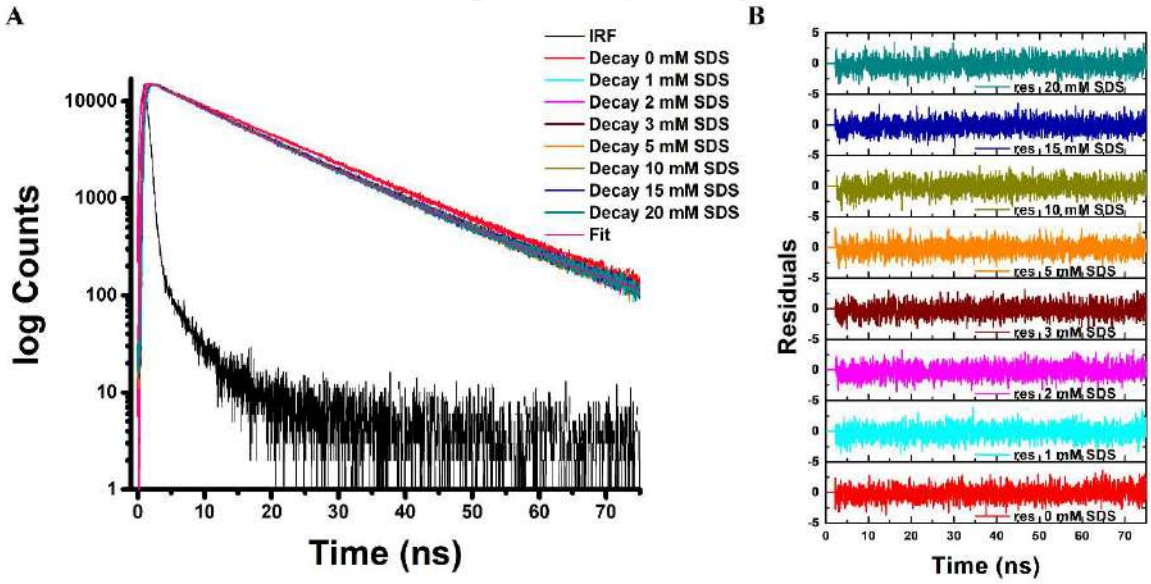


Figure 6.16: Effect of SDS on the dansyl labelled on Cys^{62} of DHN1 CW1. (**A**) Fitted time-resolved fluorescence intensity decay profile of 10 µM dansyl labeled DHN1 CW1 presence of SDS (1−20 mM) excited at λ_{ex} 340 nm, and (**B**) residuals for the fit is shown.

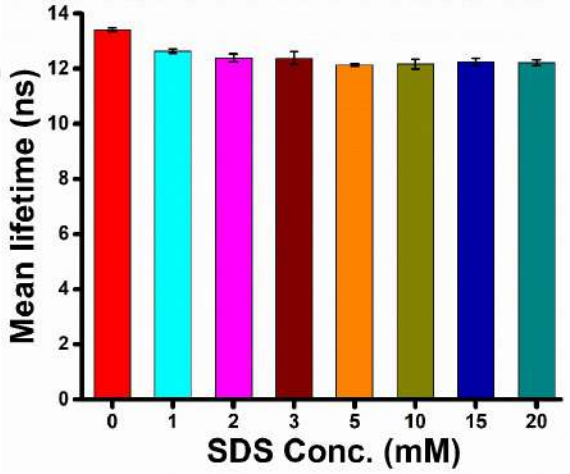


Figure: 6.17: Mean lifetime of 10 µM dansyl labelled with DHN1 CW1 lifetime (λ_{ex} 340 nm).

TH-3004_146106002

Table 6.4 Parameters obtained from the two-exponential fit for the fluorescence intensity decay of dansyl labeled on cysteine[62] of DHN1 CW1 in presence of different concentration of SDS. Values in bracket (for τ_i and τ_{Mean}) represents the standard deviation observed in two or more independent experiments, while α_i values are from the best fit.

SDS Conc. (mM)	α_1	α_2	τ_1 (ns)	τ_2 (ns)	τ_m (ns)	χ^2_R
0	0.90	0.10	14.43 (0.134)	3.60 (0.176)	13.35 (0.045)	1.05
1	0.75	0.25	14.38 (0.066)	6.84 (0.399)	12.63 (0.078)	1.08
2	0.76	0.24	14.13 (0.147)	6.75 (0.231)	12.39 (0.143)	1.04
3	0.79	0.21	14.03 (0.242)	6.56 (0.678)	12.38 (0.229)	1.01
5	0.78	0.22	13.74 (0.265)	5.83 (0.267)	12.13 (0.051)	1.04
10	0.80	0.20	13.88 (0.307)	6.18 (0.887)	12.17 (0.175)	1.09
15	0.76	0.24	13.89 (0.383)	6.07 (0.730)	12.24 (0.140)	1.08
20	0.75	0.25	14.04 (0.256)	6.71 (0.656)	12.18 (0.148)	1.05

TH-3004_146106002

6.2.5 Cryoprotection activity of DHN1:

In order to decipher the cryoprotection activity of DHN1, Lactate dehydrogenase (LDH) activity assay was performed as reported in previous studies with slight modification[175–177]. The cryoprotection activity of DHN1, BSA, and HEWL were measured under freeze/thaw stress. **Figure 6.18 A and B** illustrate the cryoprotection activity of DHN1, BSA, and HEWL under freeze/thaw stress. The enzyme activity of LDH was completely diminished when treated alone with multiple (10 times) freeze/thaw cycle whereas the activity is more than 60% (in 1:1 molar ratio) were retained in presence of the DHN1. LDH activity further increases with increase in the concentration of DHN1 as shown in **Figure 6.18 B** Similar but not identical cryoprotective trend was observed in BSA, DHN1, DHN1 CW1 and DHN1 W3. At 1:1 molar ratio DHN1 mutants are less effective than wildtype DHN1 protein as shown in **Figure 6.18 B.** More than 100% LDH activity protection was seen at high concentration of proteins. HEWL, which was used as a poor cryoprotectant displayed weak protection activity even when present in high molar concentration as reported in previous reports [175]. **Figure 6.19** depicts the LDH activity in presence of DHN1, BSA, and HEWL without freeze/thaw treatment shows a marginal increase in the activity. Our experiments are in close agreement with the previous studies showing that the K-segment (amphipathic α-helix), as well as the presence of charged amino acid has a very crucial role in LDH protection[111,232,244].

TH-3004_146106002

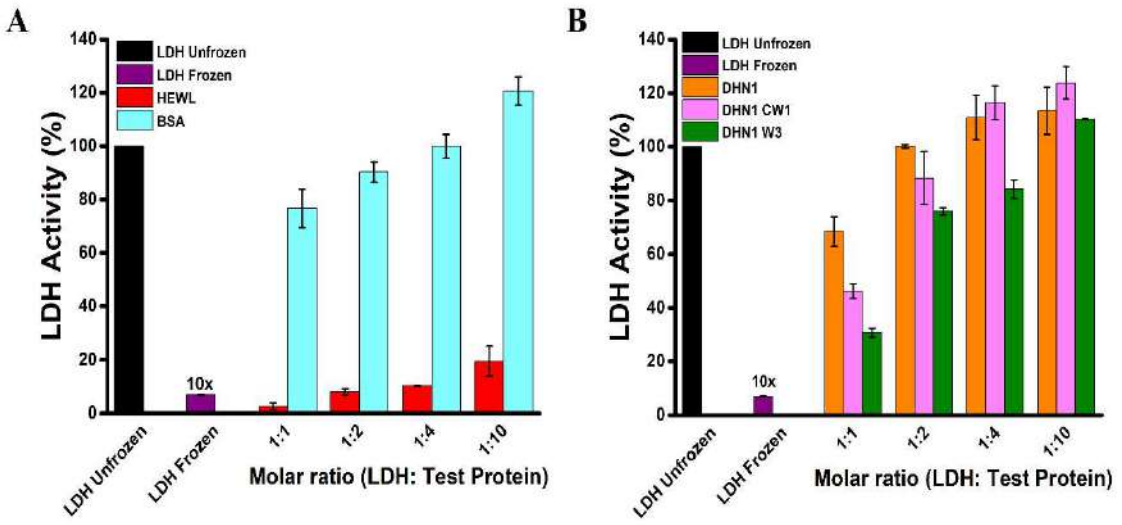


Figure 6.18: Cryoprotection activity of DHN1. LDH activity measured **(A)** in presence of BSA (positive control) and HEWL (poor cryoprotectant); and **(B)** in presence of DHN1, DHN1 CW1; and DHN1 W3. All LDH activities were assayed at 25°C and expressed as a percentage LDH activity. Activity of unfrozen LDH is taken as controls to calculate the % activity of other sample containing the same concentration of LDH. Values displayed of LDH frozen is multiplied by 10. Standard deviation shown for three individual experiments.

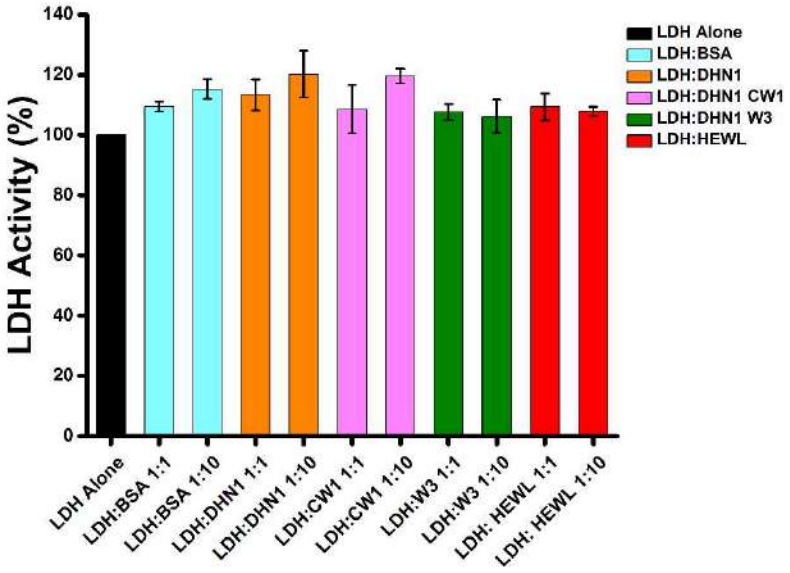


Figure 6.19: Cryoprotection activity of DHN1. LDH activity measured in presence DHN1, DHN1 CW1, DHN1 W3, BSA and HEWL proteins without freeze/thaw treatment. All the LDH activities were assayed at 25 °C and expressed as a percentage LDH activity. Standard deviation shown for three individual experiments.

TH-3004_146106002

6.2.6 Heat Protection activity of DHN1

In order to investigate the heat protection activity of DHN1, an activity assay of enzyme LDH was performed. **Figure 6.20 A, B, C, D and E** illustrate the heat protection activity of DHN1 and its mutant at 50 °C (Cyan) and 45 °C (Orange). LDH activity was reduced by ~50% and ~25% upon heating for 70 minutes at 50 °C and 45 °C, respectively. At 45 °C, ≥95% of LDH activity was retained in the presence of DHN1 and BSA. DHN1 CW1 and BSA were able to protect 100% LDH activity when present in a 1:1 molar ratio, whereas LDH activity was slightly less in presence of DHN1 and DHN1 W3. At 50 °C, ≥60% of LDH activity was retained in presence of DHN1 and BSA. DHN1 CW1 and DHN1 W3 in **Figure 6.20 B and C** shows the higher protection activity with increase in the molar ratio. **Figure 6.20 D** depicts that the heat protection activity decreased with increase in the BSA concentration at 50 °C. Interestingly, the heat protection activity in DHN1 mutants were increased compared to DHN1.

This result suggests that the DHN1 protein has an important role in stress tolerance like drought in plants. Our results in close agreement with previous studies made on other IDPs[52,109].

TH-3004_146106002

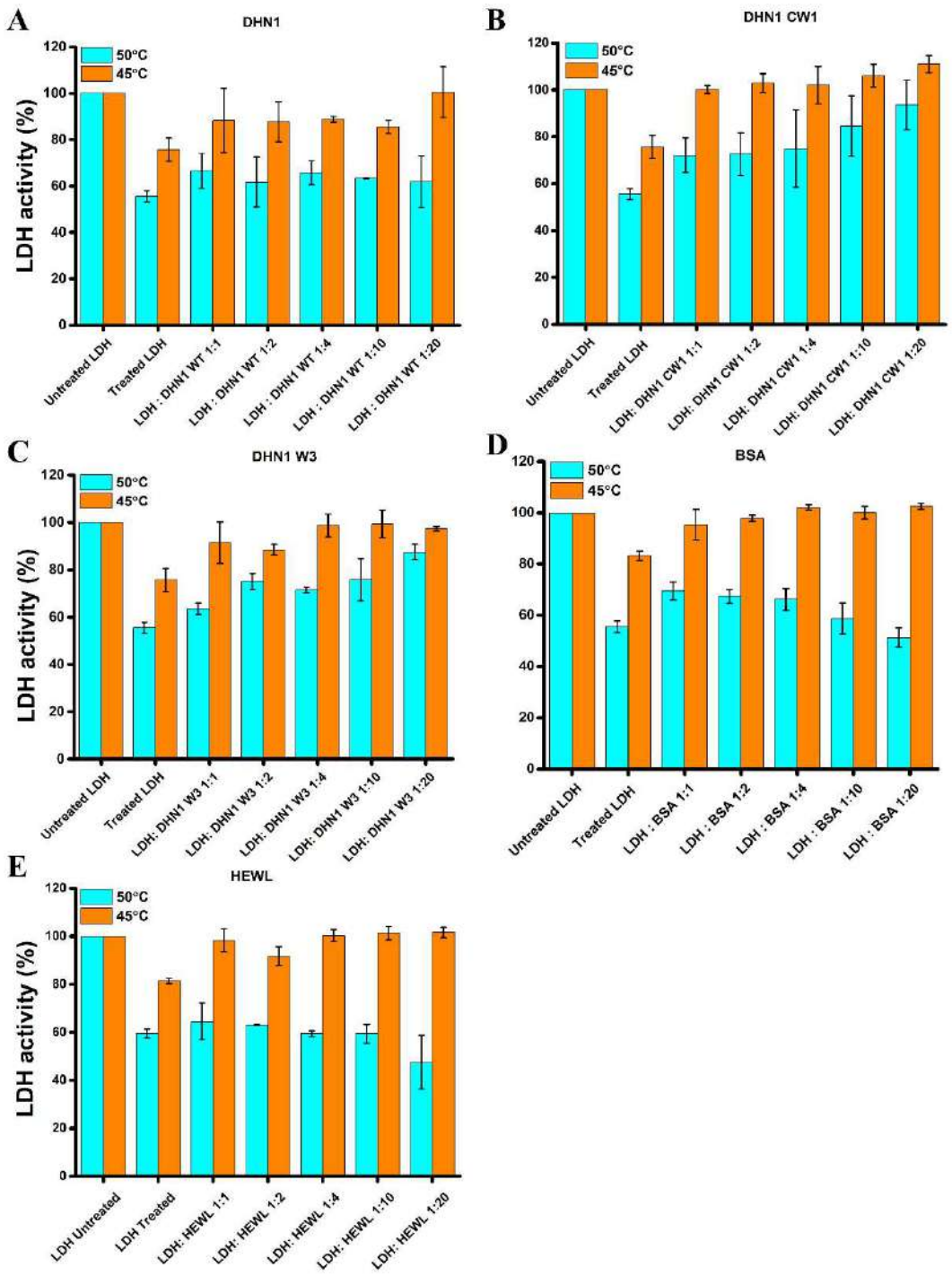


Figure 6.20: Heat-Protection activity of DHN1. LDH activity measured in presence of (**A**) DHN1; (**B**) DHN1 CW1; (**C**) DHN1 W3 (**D**)BSA; and (**E**) HEWL at 45 °C and 50 °C temperature. Samples were kept in the water bath for 70 minutes for heating at both temperatures. Standard deviation shown for three individual experiments.

TH-3004_146106002

6.2.7 Binding of DHN1 to LDH

To investigate the binding of DHN1 to LDH, steady-state fluorescence anisotropy (r_{ss}) of dansyl probe covalently labelled on single cysteine[62] of DHN1 CW1 (in excess molar ratio) was recorded in the presence of LDH. **Figure 6.21** shows steady-state fluorescence anisotropy of dansyl labeled on cysteine[62] of DHN1 CW1 in presence of LDH. A significant increase in anisotropy was observed in the presence of LDH in the 1:1, 1:2, and 1:4 molar ratios, suggesting that DHN1 binds to LDH. This binding of DHN1 to LDH (weak electrostatic) could play an important role in the cryoprotection mechanism under stressful conditions. Increase in r_{ss} with the increase in LDH concentration was observed up to 1:4 (CW1: LDH) molar ratio, which suggests towards the weak electrostatic interaction between DHN1 and LDH. This nonspecific electrostatic interaction could be attributed to the functional role of dehydrin in different stress mitigation in plants.

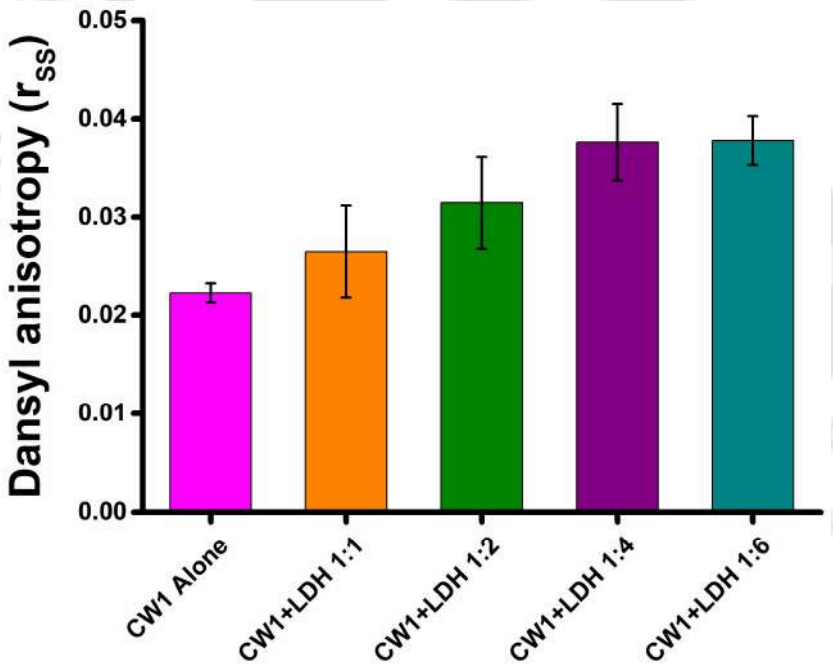


Figure 6.21: Binding of DHN1 to LDH. Steady state anisotropy of dansyl labeled on cysteine[62] of DHN1 CW1 in presence of LDH at different molar ratio. Excitation was done at 340 nm and emission at 503 nm with slit width kept for excitation and emission is 1 nm and 5 nm, respectively. The results are average of 3 measurements with the error bars showing the standard deviation.

TH-3004_146106002

6.3 Conclusions:

- DHN1 and its mutants DHN1 CW1 and DHN1 W3 are stable at higher temperature and displays the completely reversible structural transitions.

- DHN1 and its mutants gain the α -helical structure as increase in TFE.

- Presence of lower concentration of SDS initiate the structural transitions in DHN1 and its mutant proteins.

- FRET investigation in presence of SDS reveals that the two conserved K segments in DHN1 come closer to each other.

- The fluorescence studies of DHN1 CW1 and DHN1 W3 tryptophan reveals the absence of any local structure at N-terminal and C-terminal of DHN1 protein.

- DHN1 and its mutants shows the cryoprotection effect suggests the importance of DHN1 in stress tolerance mechanism.

- Heat-protection activity of DHN1 and its mutant divulges about DHN1's stability at higher temperature and significant role in heat protection of crucial enzyme in heat stress condition.

- The binding of DHN1 to LDH shows the weak interaction between LDH and DHN1 may play an important role in heat and cryoprotection activity.

TH-3004_146106002

Chapter 7

Summary and future perspectives

TH-3004_146106002

7.1 Summary:

This thesis presents a systematic investigation of an intrinsically disordered protein, DHN1, of *Zea mays*. This protein is rich in charged amino acids and critical for performing the important functions in plants during stress tolerance.

We have predicted the highly disordered structure of DHN1 protein using different prediction tools and further confirmed the characteristic disorder with various experiments. The structural analysis of DHN1 protein divulges its complete random coil conformation in the native state. Further, change in the structural transitions was studied through a newly discovered tool: ProCharTS absorption. This novel technique paved the way to investigate the IDPs in a label-free approach. However, other methods like CD and fluorescence with ProCharTS can be used for further detailed structural analysis of intrinsically disordered protein confirmations. Spectroscopic studies reveal that DHN1 protein is highly dynamic in nature and able to gain the structure close to 70 % in the presence of the binding partner.

Next, we investigated the luminescence characteristic of this novel ProCharTS in the DHN1 protein. This study suggests that ProCharTS in DHN1 protein is luminescent in nature. We found that the observed luminescence is excitation wavelength dependent and possess a significant presence in the UV-Visible region. ProCharTS luminescence has a low quantum yield and lower luminescence lifetime. The origin behind this luminescence could be charge recombination. Further, ProCharTS luminescence found to modulate the fluorescence of other chromophores like Trp in the same spectral region.

The FRET studies on DHN1 CW1 hints towards the significant role of the K-segment in binding to the membrane. Similarly, the fluorescence and CD investigation of DHN1 W3 mutant suggest that the N-terminal of DHN1 protein also take part in folding, and this part folds later in comparison to the C-terminal of DHN1 proteins. Further it can be concluded from our observation that the two K-segment located towards the C-terminal of DHN1 protein initiates the structural gain in DHN1 protein which is followed by the C-terminal region.

DHN1 protein has shown a significant cryoprotective role on LDH enzyme, suggesting that this protein plays a vital role during the plant's chilling or water desiccation state. The

TH-3004_146106002

mutation done in between K-segment and at the N-terminal of DHN1 protein has significant effect on the cryoprotection activity of DHN1.

In contrast to the cryoprotection function, the mutation between two K-segment and at the N-terminal of DHN1 protein surprisingly led to increased heat resistance or heat protection activity. The DHN1 protein has shown the weak electrostatic binding to the LDH enzyme indicating a potential mechanism for the protective function of DHN1 in plants during stressed conditions.

7.2 Future Perspectives:

The mechanism of the functions of dehydrin is not understood thoroughly. It is very important to understand the process of adaptation of plants to various abiotic stress. Detailed investigation can be made to probe the functional mechanism of DHN1 in the protection of membrane and enzymes during stressed conditions.

Further, the intrinsically disordered protein rich in charged amino acids can be investigated using the ProCharTS to gain the further insights on important aspects of this new tool and its usefulness in structural dynamics and other related issue in IDPs.

TH-3004_146106002

Appendix

TH-3004_146106002

Components of SDS-PAGE

Components for 15% (Acrylamide) Resolving Gel

Solutions	Resolving Gel (10 mL)	Stacking Gel (5 mL)
Water	2.3 mL	3.4 mL
30% Acrylamide	5 mL	830 µL
1.5 M tris(hydroxymethyl)aminomethane	2.5 mL (pH 8.8)	630 µL (pH 6.8)
10% SDS	100 µL	50 µL
10% Ammonium persulphate	100 µL	50 µL
TEMED	12 µL	8 µL

Gel Loading Buffer Composition

Gel loading buffer was prepared by mixing 0.05% (w/v) bromophenol blue, 40% (w/v) sucrose, 0.5%(w/v) SDS and 0.1 M EDTA, pH 8.0.

Staining Solution

Staining solution was prepared by dissolving 0.025% (w/v) Coomassie Blue R-250 in a solution of 45% methanol, 45% deionized water and 10% glacial acetic acid.

Destaining Solution

Destaining solution was prepared by mixing, 30% of methanol, 10% of glacial acetic acid and 60% of deionized water.

TH-3004_146106002

Solutions for Lowry method

Solution A (2% Sodium carbonate in 0.1% NaOH), 100 mL

NaOH	0.47 g
Na_2CO_3	2 g

Solution B (2.37% Potassium sodium tartrate in water), 50 mL

$KNaC_4H_4O_6.4H_2O$	1.185 g

Solution C (1.56% Copper sulfate in water), 50 mL

$CuSO_4.5H_2O$	0.78 g

Reagent I: Preparation by mixing 48 mL of solution A; 1 mL of Solution B and 1 mL of Solution C

Reagent II: It was prepared by mixing of equal volume of Folin's reagent and water (Stored in dark)

Model used for the fitting of the change in ellipticity at 222 nm and 198 nm as function of SDS concentrations.

* Exponential Decay Fit, $y = A1*\exp(-x/xt)+y0$

\# Logistic Fit, $y = A2 + (A1-A2)/(1+(x/x0)^p)$

Table A1: Model used for the fitting the change in tryptophan anisotropy as function of SDS concentration

Protein	Non-linear curve fit (Hill1)* for change in Trp anisotropy		
	k	n	R^2
DHN1 CW1	0.16 ± 0.03	3.65 ± 2.30	0.94
DHN1 W3	0.54 ± 0.02	5.20 ± 1.3	0.99

***Hill1,** $y = Start + (End-Start)*x^n/(k^n + x^n)$

TH-3004_146106002

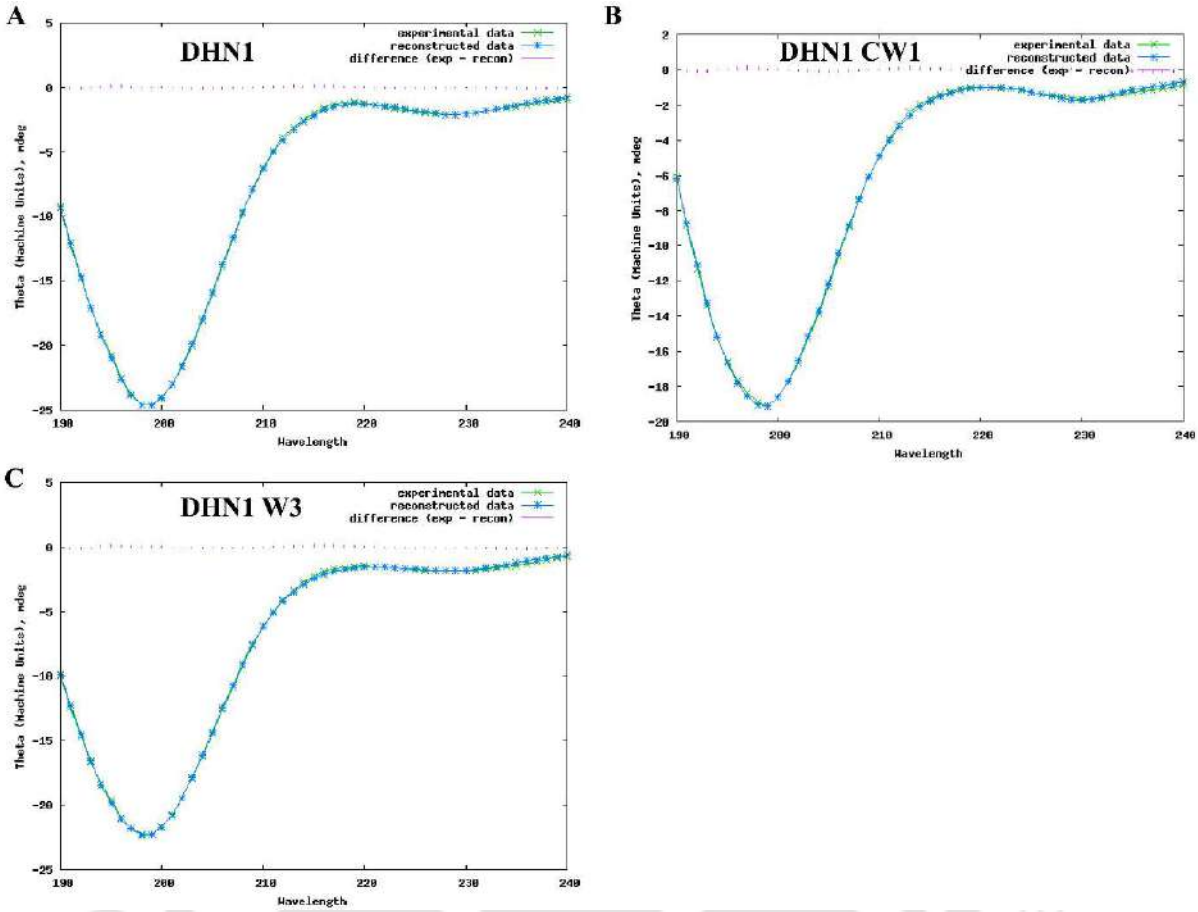


Figure A1: Fitted CD spectra of DHN1, DHN1 CW1 and DHN1 W3 dissolved in water.

TH-3004_146106002

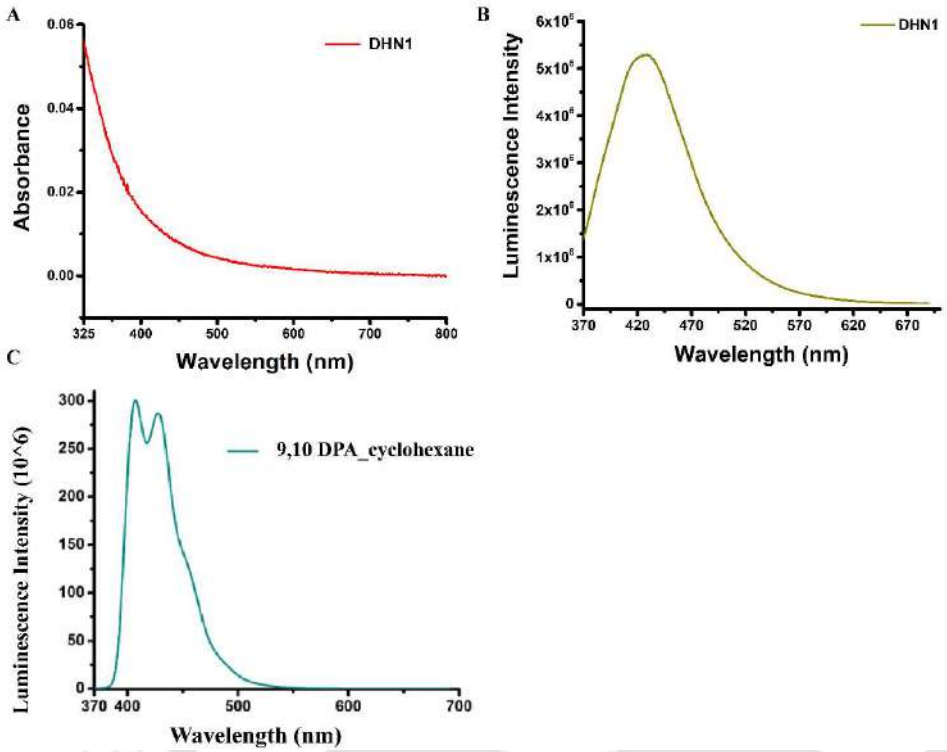


Figure A2: Absorption and emission spectra used for quantum yield calculation. Luminescence spectra was collected for λ_{ex} 355 nm between 370-690 nm using excitation slit width of 1 nm and emission slit width of 5 nm.

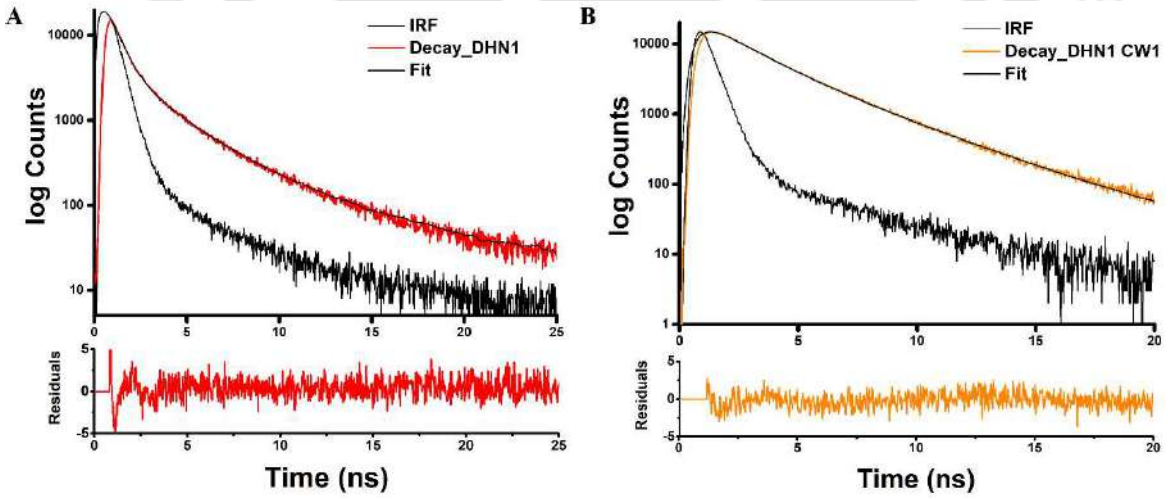


Figure A3: Time resolved fluorescence intensity decay of DHN1 and DHN1 CW1. Fitted time-resolved fluorescence intensity decay fit for the **(A)** DHN1, **(B)** DHN1 CW1 excited at λ_{ex} 295 nm. Emission was collected using the bandpass filter (340 ±20 nm). Residuals for the fit is shown in panel below.

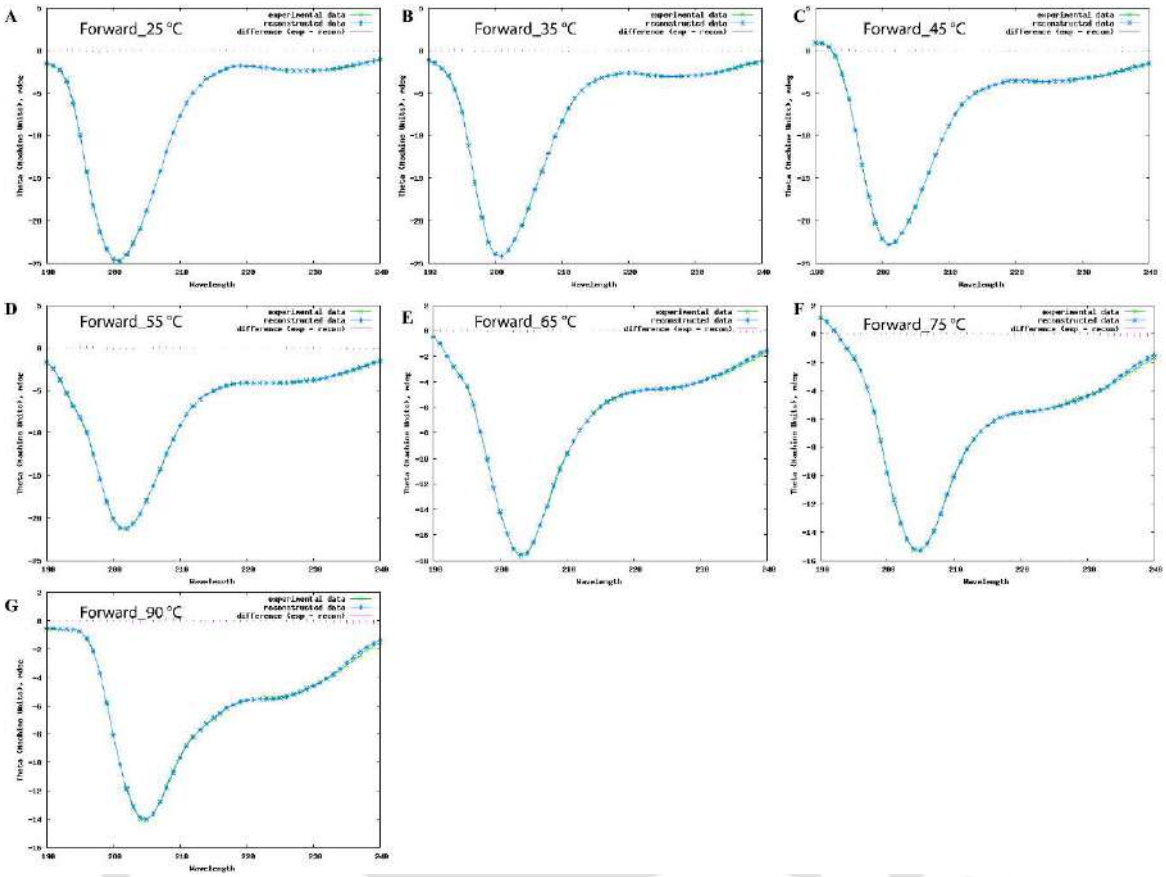


Figure A4: Fitted CD Spectra of DHN1 at high temperature by using DichroWeb server. The fitted data shown here is only for heating DHN1 from 25 to 90 °C.

TH-3004_146106002

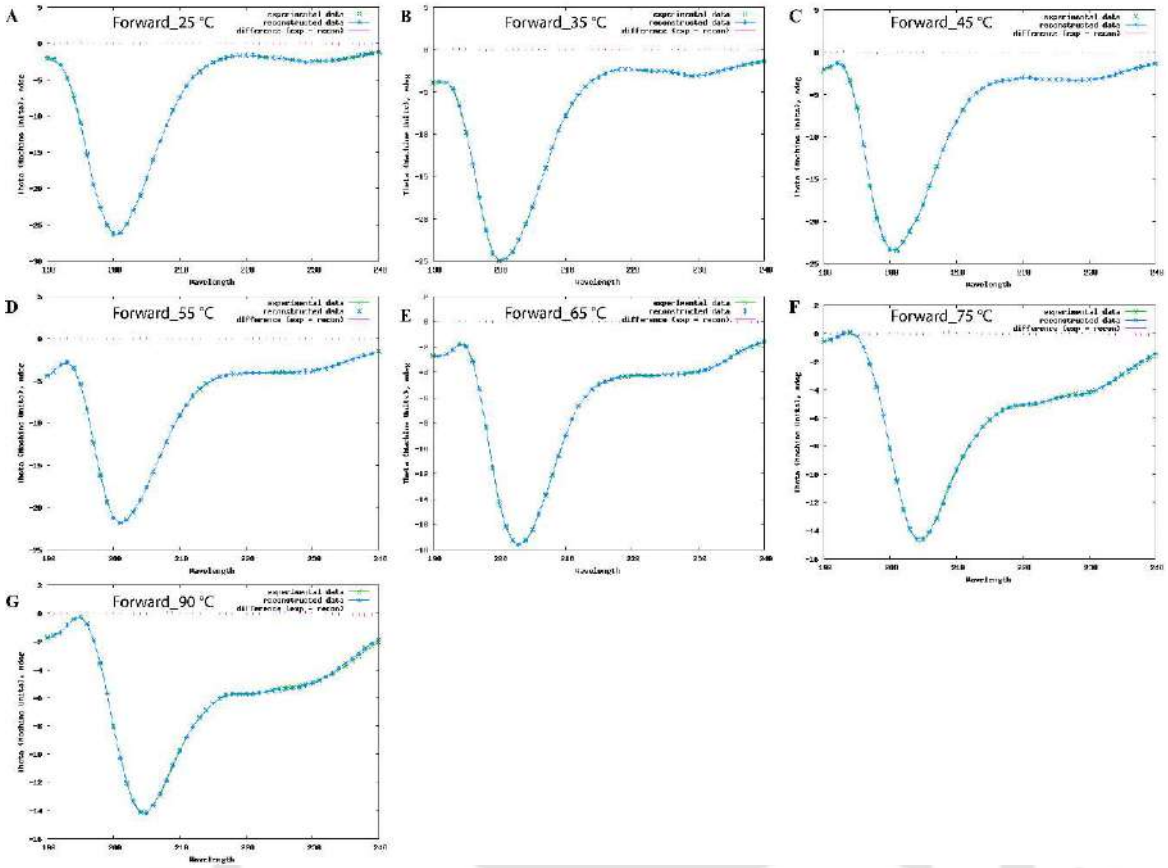


Figure A5: Fitted CD Spectra of DHN1 CW1 at high temperature by using DichroWeb server. The fitted data shown here is only for heating DHN1 CW1 from 25 to 90 °C.

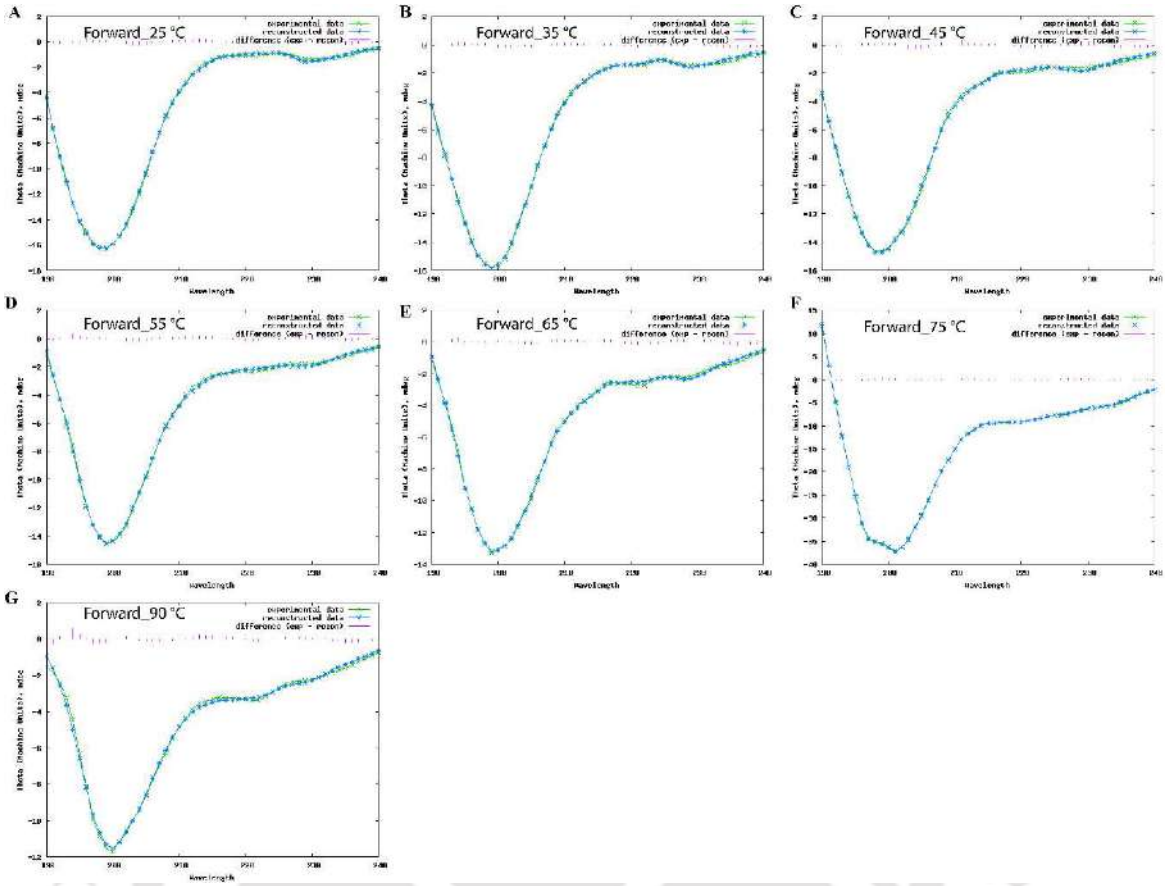


Figure A6: Fitted CD Spectra of DHN1 W3 at high temperature by using DichroWeb.

The fitted data shown here is only for heating DHN1 W3 from 25 to 90 °C.

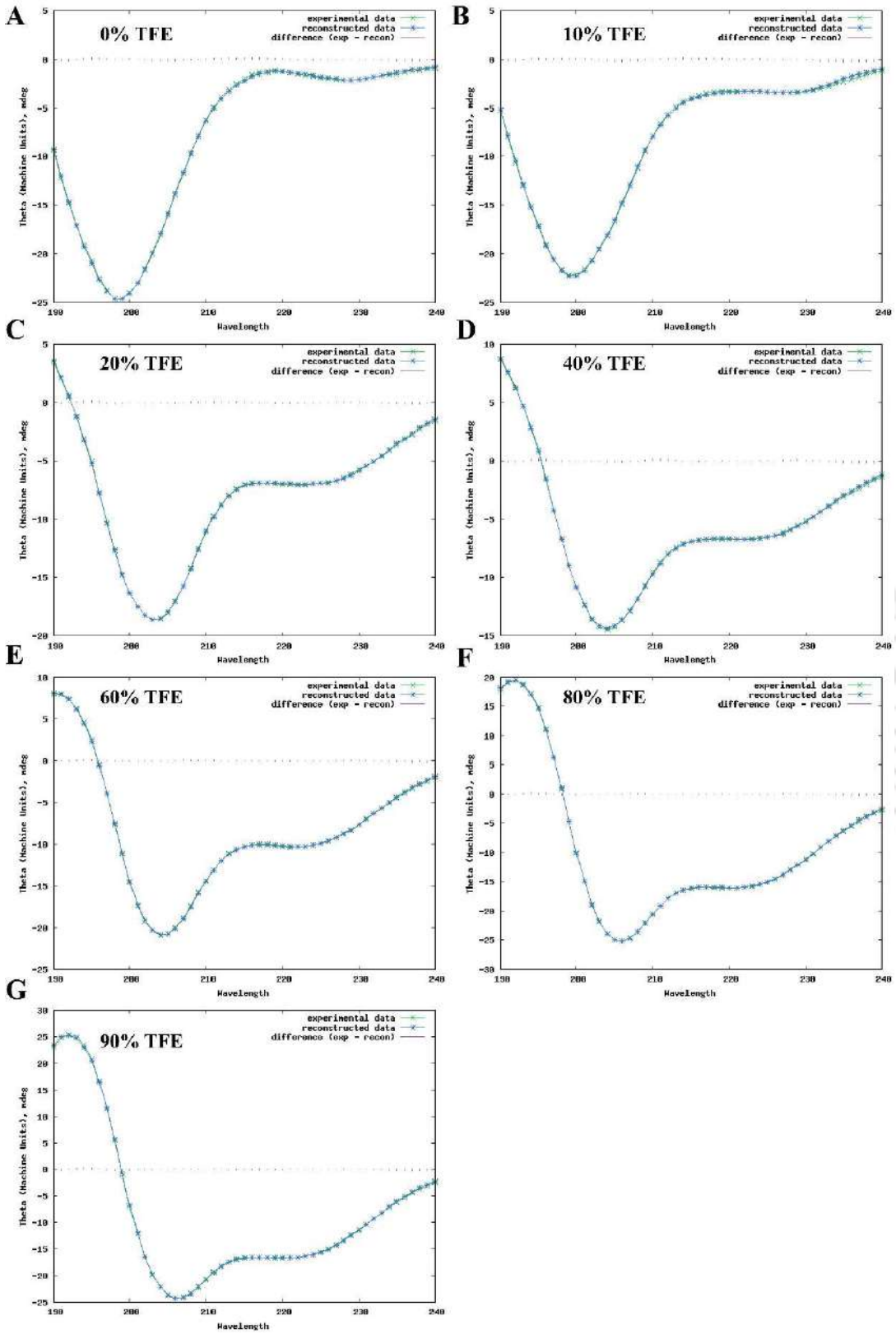


Figure A7: Fitted CD Spectra of DHN1 at various concentration of TFE (10−90%) by using DichroWeb server.

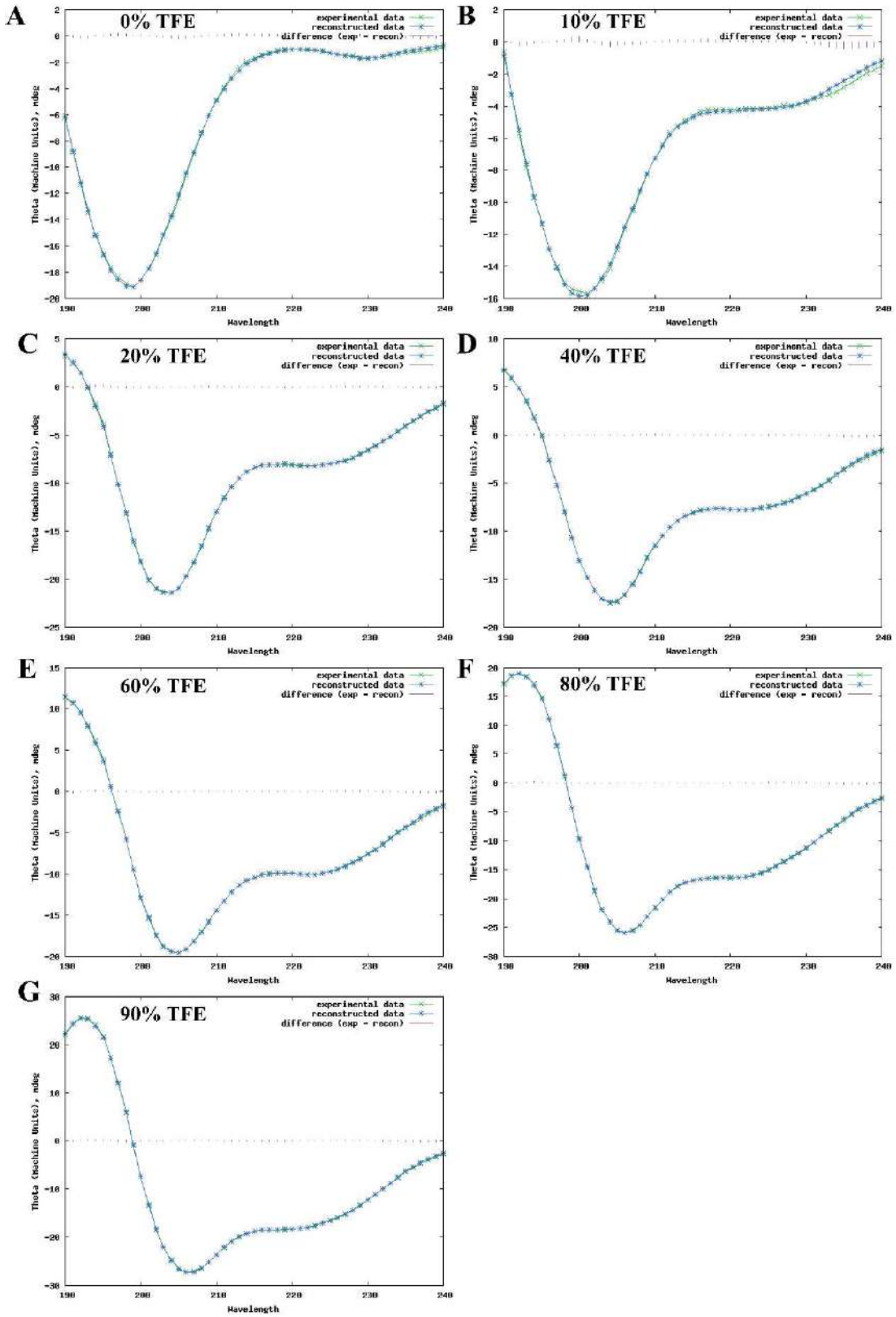


Figure A8: Fitted CD Spectra of DHN1 CW1 at various concentration of TFE (10−90%) by using DichroWeb server.

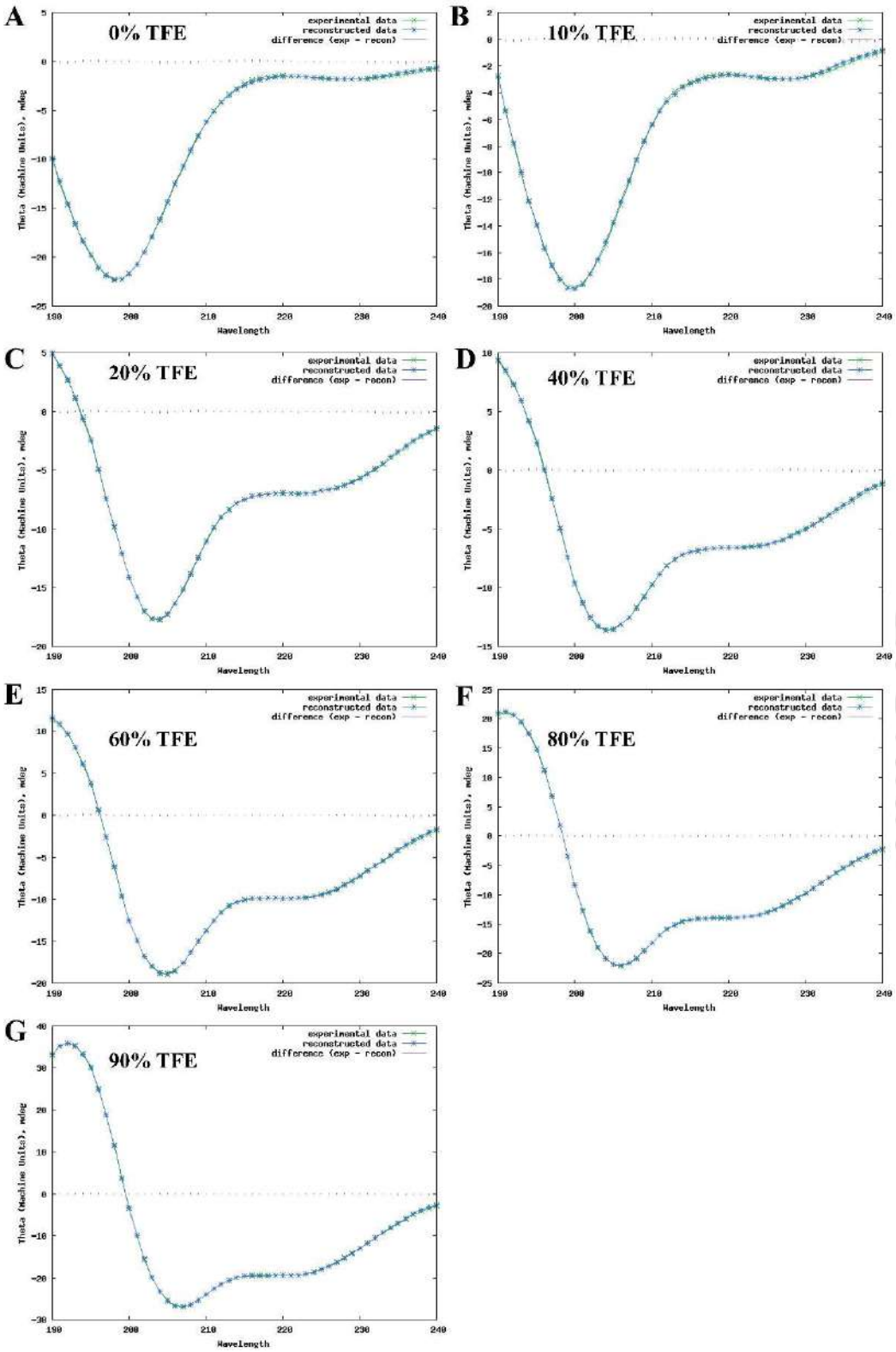


Figure A9: Fitted CD Spectra of DHN1 W3 at various concentration of TFE (10−90%) by using DichroWeb server.

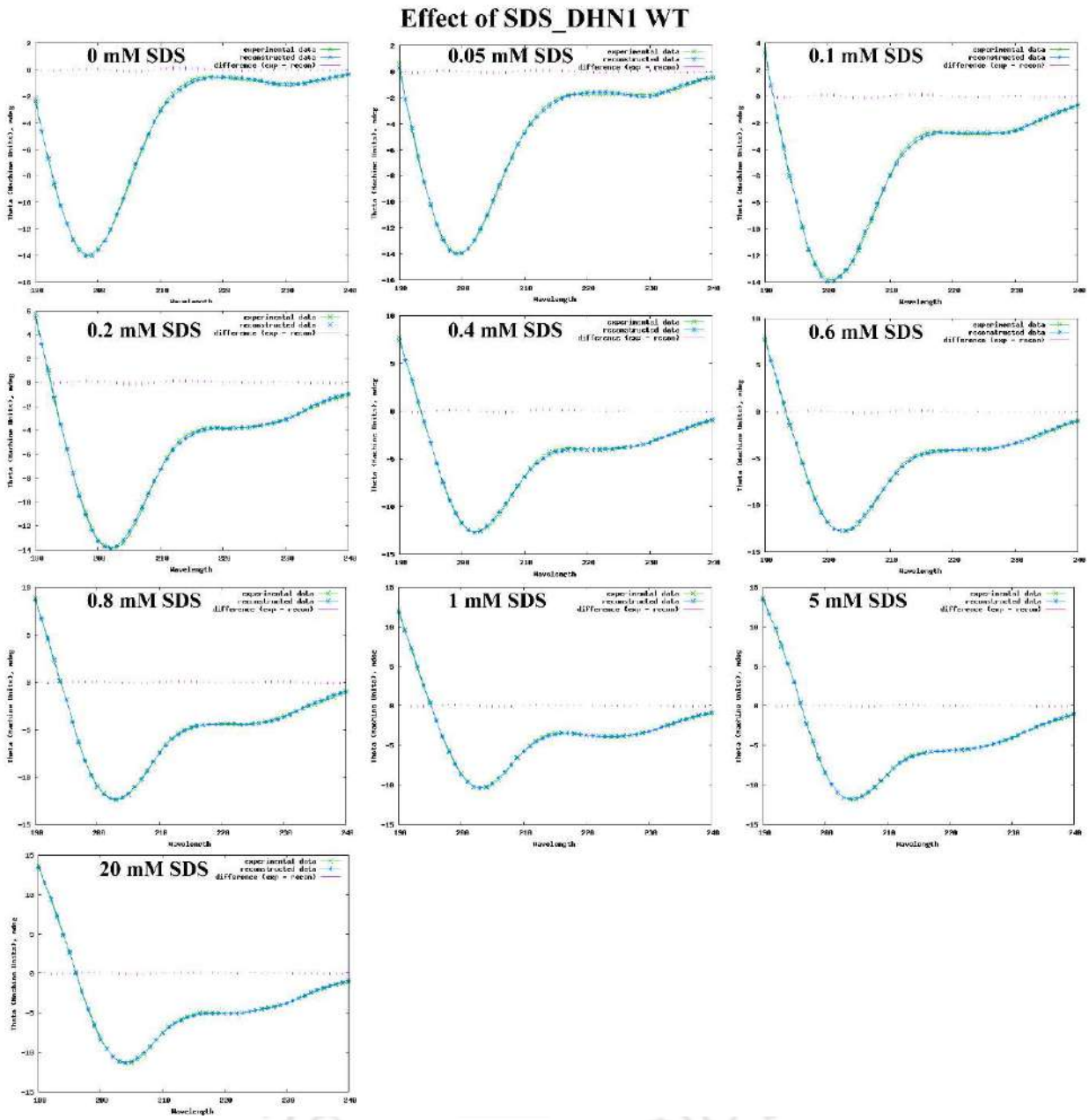


Figure A10: Fitted CD Spectra of DHN1 in presence of different concentration of SDS (0.05−20 mM) by using DichroWeb server.

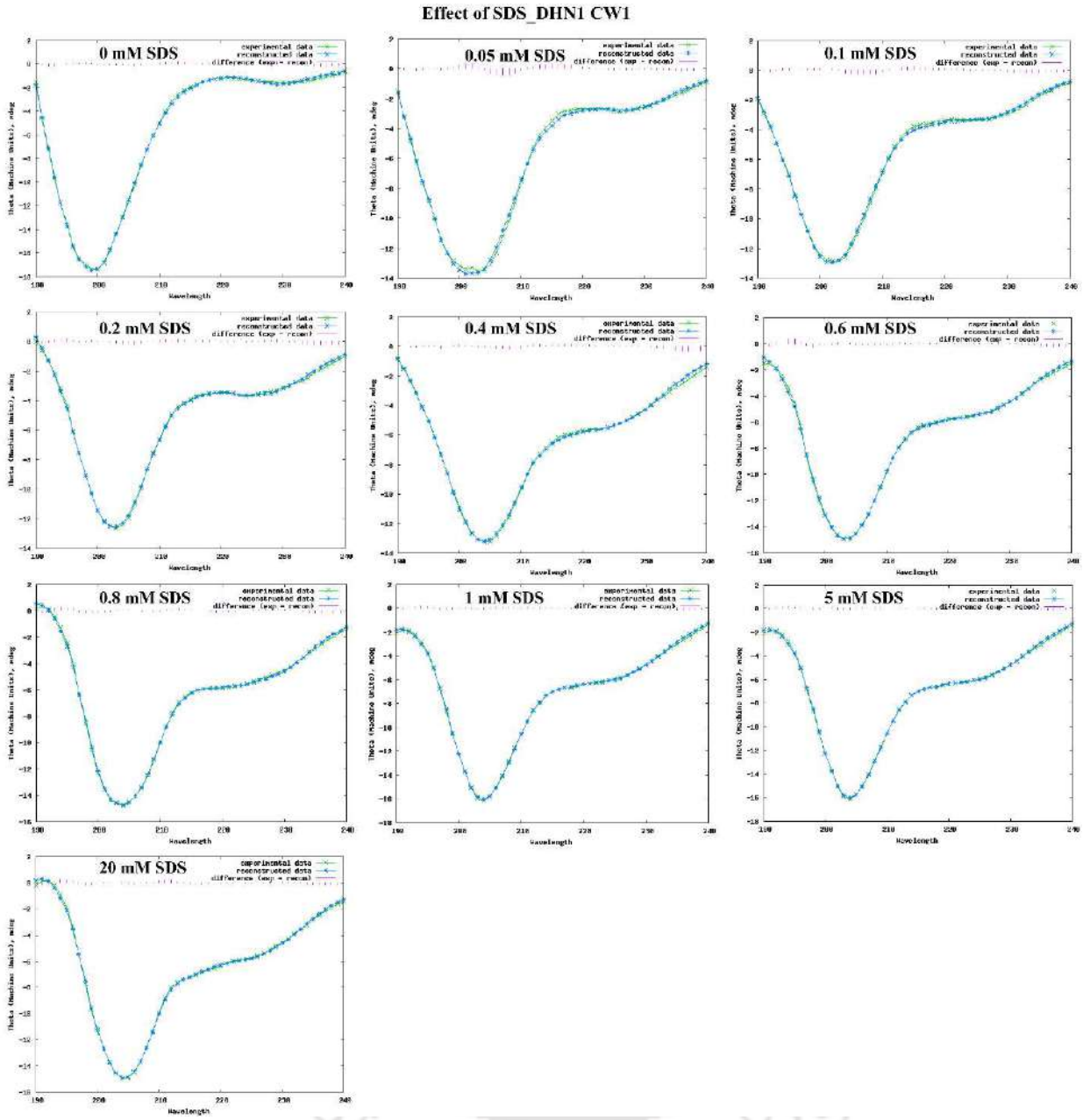


Figure A11: Fitted CD Spectra of DHN1 CW1 in presence of different concentration of SDS (0.05−20 mM) by using DichroWeb server.

TH-3004_146106002

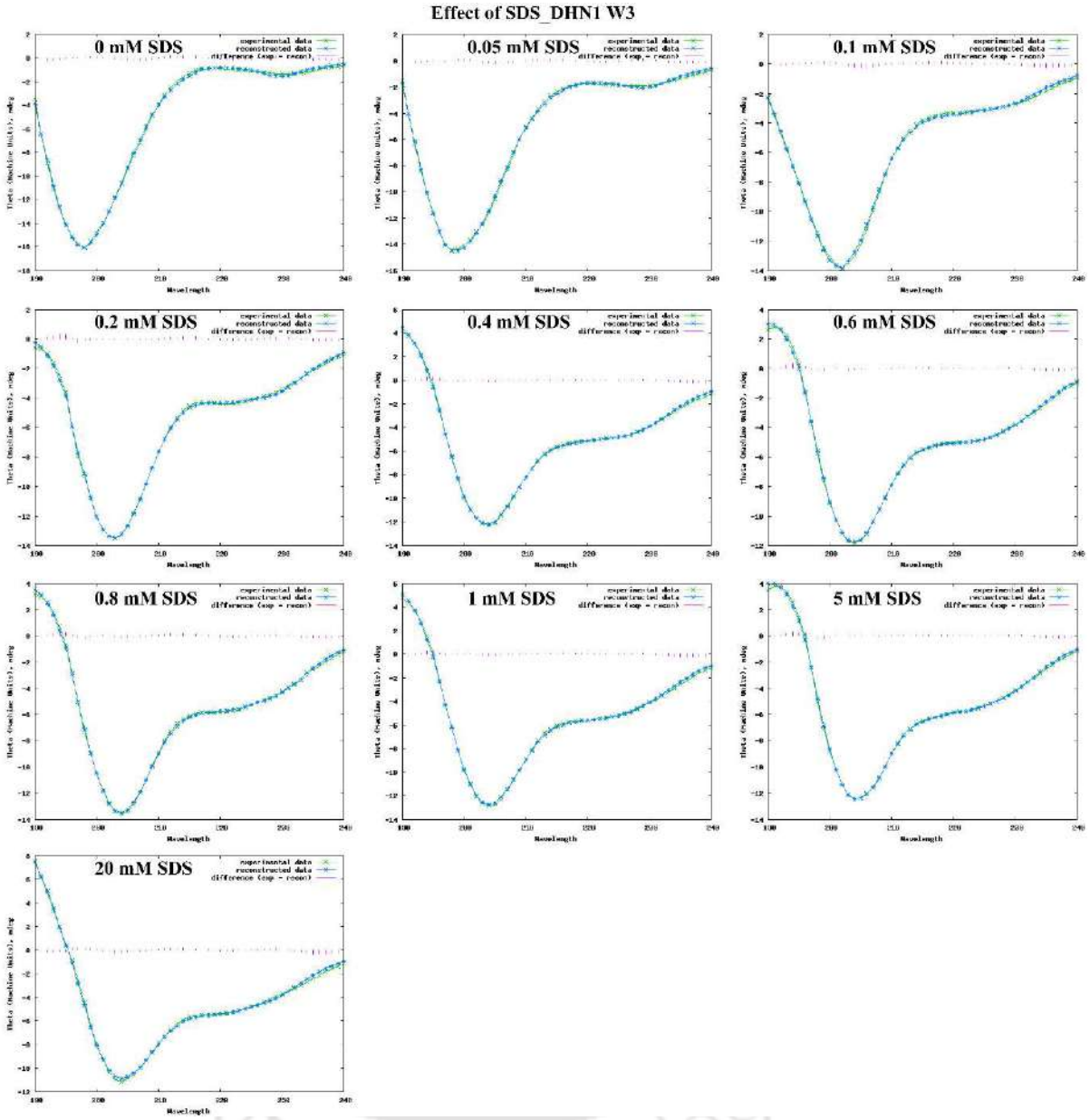


Figure A12: Fitted CD Spectra of DHN1 W3 in presence of different concentration of SDS (0.05−20 mM) by using DichroWeb server.

TH-3004_146106002

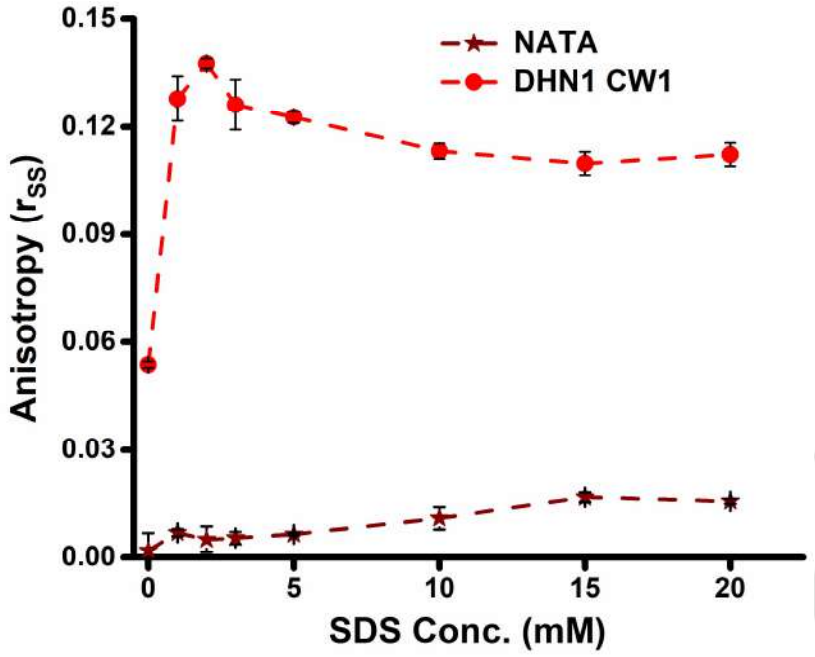


Figure A13 Steady state fluorescence anisotropy of DHN1 CW1 and NATA in presence of SDS. Fluorescence anisotropy DHN1 and NATA (20 μM) in different concentration of SDS (1-20 mM) upon excitation at 295 nm. The anisotropy was collected at emission maxima of 350 nm. Slit width kept for excitation and emission is 1 and 5 nm, respectively.

TH-3004_146106002

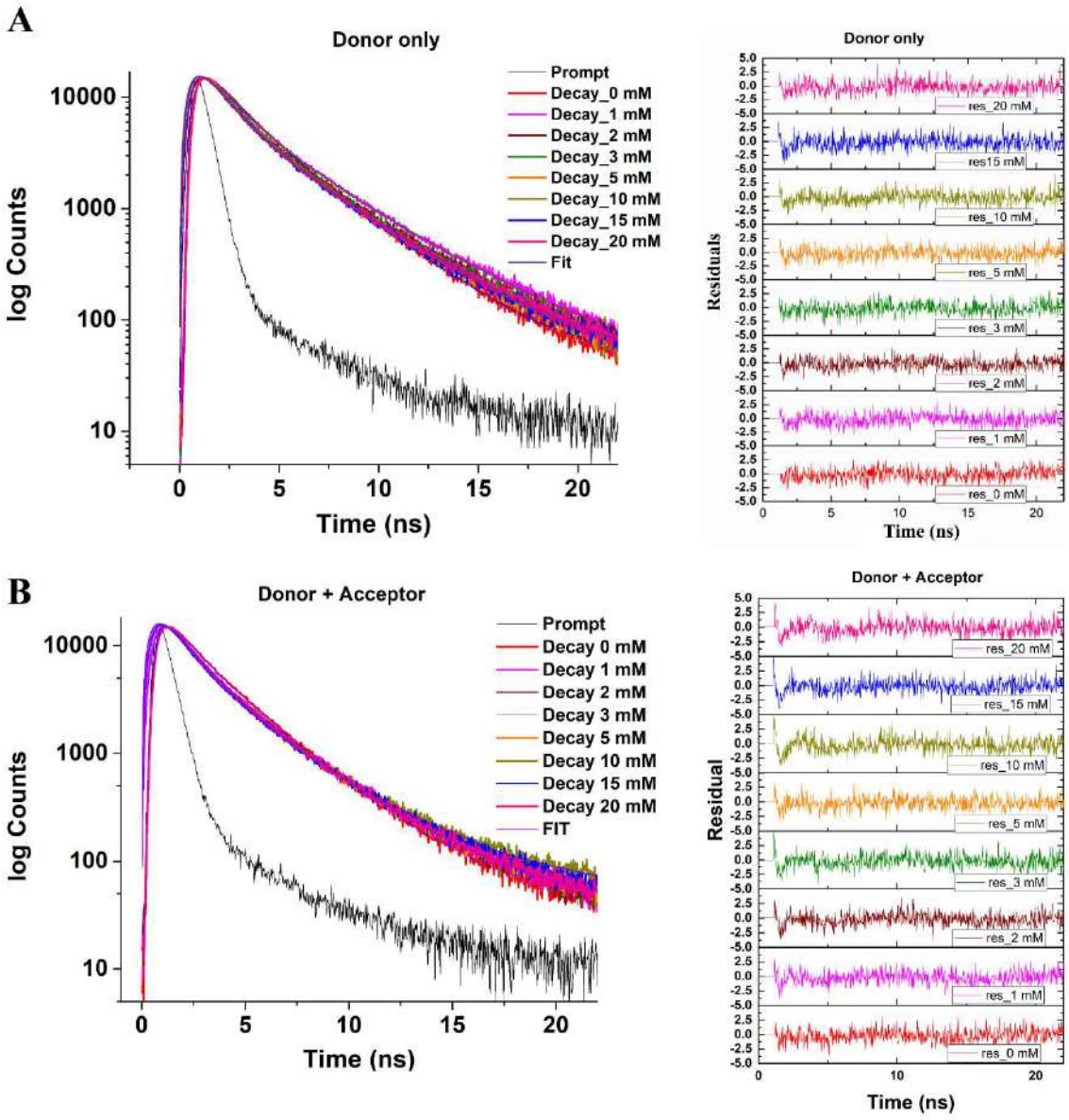


Figure A14 Effect of SDS on the Distance between Cys^{62} (dansyl)- Trp^{122}(indole) in DHN1 CW1. Fitted time-resolved fluorescence intensity decay fit for the 10 µM DHN1 CW1 tryptophan (donor) excited at λ_{ex} of 295 nm in presence of **(A)** donor only **(B)** donor+ acceptor, at different concentration of SDS (1-20 mM). Residuals for the fit is shown in right panel.

TH-3004_146106002

LIST OF PUBLICATIONS AND CONFERENCES

TH-3004_146106002

List of Publications

1. Ansari MZ, Kumar A, **Ahari D**, Priyadarshi A, Lolla P, Bhandari R, et al. Protein charge transfer absorption spectra: An intrinsic probe to monitor structural and oligomeric transitions in proteins. Faraday Discuss. 2018;207:91–113.

2. Kumar A, **Ahari D**, Priyadarshi A, Ziauddin Ansari M, Swaminathan R. Weak Intrinsic Luminescence in Monomeric Proteins Arising from Charge Recombination. J Phys Chem B. 2020;124(14):2731–2746.

3. Kumar A, Alom SE, **Ahari D**, Priyadarshi A, Ansari MZ, Swaminathan R. Role of Charged Amino Acids in Sullying the Fluorescence of Tryptophan or Conjugated Dansyl Probe in Monomeric Proteins. Biochemistry. 2022;61(5):339–353.

4. **Ahari D,** Sahil K, Rangan L, Swaminathan R. Investigation of structural and functional role of Intrinsically Disordered Protein, dehydrin and its mutants. [**Manuscript under preparation**]

List of conference abstract publications:

1. **Ahari D**, Sahil K, Rangan L, Swaminathan R. SDS induced transition and cryoprotective function of plant IDP, dehydrin and its mutant. Biophys J. 2022; 121 (3, Supplement 1):201a

2. Kumar A, **Ahari D**, Priyadarshi A, Ansari MZ, Swaminathan R. Weak Intrinsic Luminescence in monomeric Proteins Arising from Charge Recombination. Biophys J. 2020; 118 (3, Supplement 1):468a

3. Kumar A, Alom SE, Priyadarshi A, **Ahari D**, Ansari MZ, Swaminathan R. Unraveling the Origin of Multi-Exponential Fluorescence Intensity Decay of Tryptophan in Proteins. Biophys J. 2020;118(3, Supplement 1):469a.

4. Ansari, M. Z.; Kumar, A.; **Ahari, D.**; Priyadarshi, A.; Lolla, P.; Bhandari, R.; Swaminathan, R. 2018. Protein Charge Transfer Absorption Spectra: An Intrinsic Probe to Monitor Structural and Oligomeric Transitions in Proteins. Biophysical Journal 114: 586a

List of conferences attended

1. 66th Biophysical Society Annual Meeting (BPS2022), San Francisco, California (February 23, **2022**) – Online poster presentation

2. 64th Biophysical Society Annual Meeting (BPS2020), San Diego, California (February 15-19, **2020**) -Poster presentation

3. National Workshop on Fluorescence and Raman Spectroscopy, IIT Bombay-TIFR Mumbai (December, **2020**) – online poster presentation

4. Research Conclave, IIT Guwahati (March 14-17, **2019**) – Poster presentation

TH-3004_146106002

5. National Workshop on Fluorescence and Raman Spectroscopy, JNU-IIT Delhi-TIFR Mumbai (Nov 12-17, **2018) –** Poster presentation

6. Research Conclave, IIT Guwahati (March 8-11, **2018**) – Poster presentation

7. National Workshop on Fluorescence and Raman Spectroscopy, IIT Guwahati-TIFR Mumbai (December 17-21, **2017**) – Poster presentation and volunteer

8. International Conference on "Intrinsically Disordered Proteins: Forms, Function and Disease" IISER Mohali (December 9-12, **2017**) – Poster presentation

9. Horiba Optical School, JNCASR Bengaluru (May 22-26,**2017**) – Workshop and poster presentation

10. Indo-Japan Symposium on "Hope from Herbs: Research -based Care & cure Potential" IIT Guwahati (India)-AIST (Japan) (May 8-9, **2017**)

11. Third Symposium on Protein Folding and Dynamics at NCBS Bangalore (Nov 8-11, **2016**) – Poster presentation

12. Workshop on "Intellectual Property Right" IIT Guwahati (30th Nov -1st Dec**, 2016)**

13. National Conference on "Recent Development in Medical Biotechnology and Structure based Drug Designing" at IIT Guwahati (December 6-7, **2015**)

Indian Institute of Technology Guwahati

References

1. Campbell, N. A. *Biology*. (Pearson Education Australia, 2009).

2. Alberts, Bruce, Johnson, Alexandar, Julian, Lewis, Martin, Raff, Keith, Roberts, and P. W. *Molecular biology of the cell*. (Garland Publisher ©1989).

3. M. Berge, Jeremy , L.tymoczko, John , Stryer, L. Biochemistry. (2546).

4. Daniel, R. M., Dunn, R. V, Finney, J. L. & Smith, J. C. The role of dynamics in enzyme activity. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 69–92 (2003).

5. di Prisco, G., Condo, S. G., Tamburrini, M. & Giardina, B. Oxygen transport in extreme environments. *Trends Biochem. Sci.* **16**, 471–474 (1991).

6. Kendrew, J. C. *et al.* A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **181**, 662–666 (1958).

7. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R. & Klug, A. Protein disk of tobacco mosaic virus at 2.8 A resolution showing the interactions within and between subunits. *Nature* **276**, 362–368 (1978).

8. Bode, W., Schwager, P. & Huber, R. The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. The refined crystal structures of the bovine trypsinogen-pancreatic trypsin inhibitor complex and of its ternary complex with Ile-Val at 1.9 A resolution. *J. Mol. Biol.* **118**, 99–112 (1978).

9. Huber, R. & Bennett, W. S. J. Functional significance of flexibility in proteins. *Biopolymers* **22**, 261–279 (1983).

10. Holt, C. & Sawyer, L. Caseins as rheomorphic proteins: interpretation of primary and secondary structures of the αS1-{,} β- and κ-caseins. *J. Chem. Soc.{,} Faraday Trans.* **89**, 2683–2692 (1993).

11. Pullen, R. A., Jenkins, J. A., Tickle, I. J., Wood, S. P. & Blundell, T. L. The relation of polypeptide hormone structure and flexibility to receptor binding: the relevance of X-ray studies on insulins, glucagon and human placental lactogen. *Mol. Cell. Biochem.* **8**, 5–20 (1975).

12. Schweers, O., Schönbrunn-Hanebeck, E., Marx, A. & Mandelkow, E. Structural

TH-3004_146106002

studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *J. Biol. Chem.* **269**, 24290–24297 (1994).

13. Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A. & Lansbury, P. T. J. NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* **35**, 13709–13715 (1996).

14. Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are 'natively unfolded' proteins unstructured under physiologic conditions? *Proteins* **41**, 415–427 (2000).

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

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

17. Dunker, A. K. *et al.* Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**, 26–59 (2001).

18. Fuxreiter, M. *et al.* Malleable machines take shape in eukaryotic transcriptional regulation. *Nat. Chem. Biol.* **4**, 728–737 (2008).

19. Livesay, D. R. Protein dynamics: dancing on an ever-changing free energy stage. *Current opinion in pharmacology* vol. 10 706–708 (2010).

20. Babu, M. M., van der Lee, R., de Groot, N. S. & Gsponer, J. Intrinsically disordered proteins: Regulation and disease. *Curr. Opin. Struct. Biol.* **21**, 432–440 (2011).

21. Dunker, A. K. *et al.* What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disord. Proteins* **1**, e24157 (2013).

22. Uversky, V. N. Introduction to intrinsically disordered proteins (IDPs). *Chem. Rev.* **114**, 6557–6560 (2014).

23. Dunker, A. K. *et al.* The unfoldomics decade: An update on intrinsically disordered proteins. *BMC Genomics* **9**, 1–26 (2008).

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

TH-3004_146106002

25. Dunker, A. K. *et al.* Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**, 26–59 (2001).

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

27. Radivojac, P. *et al.* Intrinsic disorder and functional proteomics. *Biophys. J.* **92**, 1439–1456 (2007).

28. Bürgi, J., Xue, B., Uversky, V. N. & van der Goot, F. G. Intrinsic Disorder in Transmembrane Proteins: Roles in Signaling and Topology Prediction. *PLoS One* **11**, 1–21 (2016).

29. Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 (2002).

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

31. Fox, S. J. & Kannan, S. Probing the dynamics of disorder. *Prog. Biophys. Mol. Biol.* **128**, 57–62 (2017).

32. Radivojac, P. *et al.* Intrinsic disorder and functional proteomics. *Biophys. J.* **92**, 1439–1456 (2007).

33. Romero, P. *et al.* Sequence complexity of disordered protein. *Proteins* **42**, 38–48 (2001).

34. WILLIAMS, R. M. *et al.* THE PROTEIN NON-FOLDING PROBLEM: AMINO ACID DETERMINANTS OF INTRINSIC ORDER AND DISORDER. in *Biocomputing 2001* 89–100 .

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

36. Oldfield, C. J. *et al.* Comparing and combining predictors of mostly disordered proteins. *Biochemistry* **44**, 1989–2000 (2005).

37. Dunker, A. K., Obradovic, Z., Romero, P., Garner, E. C. & Brown, C. J. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome*

TH-3004_146106002

Inform. **11**, 161–171 (2000).

38. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* **18**, 343–384 (2005).

39. Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M. & Uversky, V. N. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **272**, 5129–5148 (2005).

40. Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z. & Dunker, A. K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **323**, 573–584 (2002).

41. Uversky, V. N. Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: Another illustration of the D 2 concept. *Expert Rev. Proteomics* **7**, 543–564 (2010).

42. Crick, S. L., Jayaraman, M., Frieden, C., Wetzel, R. & Pappu, R. V. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16764–16769 (2006).

43. Dougan, L., Li, J., Badilla, C. L., Berne, B. J. & Fernandez, J. M. Single homopolypeptide chains collapse into mechanically rigid conformations. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 12605–12610 (2009).

44. Andreas, M., Karin, J. & Thomas, K. End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. *Proc. Natl. Acad. Sci.* **103**, 12394–12399 (2006).

45. Mukhopadhyay, S., Krishnan, R., Lemke, E. A., Lindquist, S. & Deniz, A. A. A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2649–2654 (2007).

46. Vitalis, A., Wang, X. & Pappu, R. V. Quantitative characterization of intrinsic disorder in polyglutamine: insights from analysis based on polymer theories.

TH-3004_146106002

Biophys. J. **93**, 1923–1937 (2007).

47. Walters, R. H. & Murphy, R. M. Examining polyglutamine peptide length: a connection between collapsed conformations and increased aggregation. *J. Mol. Biol.* **393**, 978–992 (2009).

48. Wang, X., Vitalis, A., Wyczalkowski, M. A. & Pappu, R. V. Characterizing the conformational ensemble of monomeric polyglutamine. *Proteins* **63**, 297–311 (2006).

49. Cumberworth, A., Lamour, G., Babu, M. M. & Gsponer, J. Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. *Biochem. J.* **454**, 361–369 (2013).

50. Cordero, O. J., Sarandeses, C. S., Lopez, J. L. & Nogueira, M. On the anomalous behaviour on gel-filtration and SDS-electrophoresis of prothymosin-alpha. *Biochem. Int.* **28**, 1117–24 (1992).

51. Heyen, B. J. *et al.* The Calcium-Binding Activity of a Vacuole-Associated, Dehydrin-Like Protein Is Regulated by Phosphorylation. *Plant Physiol.* **130**, 675–687 (2002).

52. Kovacs, D., Kalmar, E., Torok, Z. & Tompa, P. Chaperone activity of ERD10 and ERD14, two disordered stress-related plant proteins. *Plant Physiol.* **147**, 381–390 (2008).

53. Dunker, A. K., Brown, C. J. & Obradovic, Z. Identification and functions of usefully disordered proteins. *Adv. Protein Chem.* **62**, 25–49 (2002).

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

55. Galea, C. A. *et al.* Proteomic studies of the intrinsically unstructured mammalian proteome. *J. Proteome Res.* **5**, 2839—2848 (2006).

56. Fontana, A. *et al.* Probing the conformational state of apomyoglobin by limited proteolysis. *J. Mol. Biol.* **266**, 223–230 (1997).

TH-3004_146106002

57. Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E. & Zambonin, M. Probing the partly folded states of proteins by limited proteolysis. *Fold. Des.* **2**, R17-26 (1997).

58. Hubbard, S. J., Eisenmenger, F. & Thornton, J. M. Modeling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci.* **3**, 757–768 (1994).

59. Wiggins, C. M. *et al.* BIMEL, an intrinsically disordered protein, is degraded by 20S proteasomes in the absence of poly-ubiquitylation. *J. Cell Sci.* **124**, 969–977 (2011).

60. Kim, T. D., Ryu, H. J., Cho, H. I., Yang, C. H. & Kim, J. Thermal behavior of proteins: heat-resistant proteins and their heat-induced secondary structural changes. *Biochemistry* **39**, 14839–14846 (2000).

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

62. Ekman, D., Light, S., Björklund, Å. K. & Elofsson, A. What properties characterize the hub proteins of the protein-protein interaction network of Saccharomyces cerevisiae? *Genome Biol.* **7**, R45 (2006).

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

64. Patil, A. & Nakamura, H. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.* **580**, 2041–2045 (2006).

65. Tompa, P. & Csermely, P. The role of structural disorder in the function of RNA and protein chaperones. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **18**, 1169–1175 (2004).

66. Cox, C. J. *et al.* The regions of securin and cyclin B proteins recognized by the ubiquitination machinery are natively unfolded. *FEBS Lett.* **527**, 303–308 (2002).

TH-3004_146106002

67. Iakoucheva, L. M. *et al.* The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049 (2004).

68. Olashaw, N., Bagui, T. K. & Pledger, W. J. Cell cycle control: a complex issue. *Cell Cycle* **3**, 263–264 (2004).

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

70. Gianni, S., Dogan, J. & Jemth, P. Coupled binding and folding of intrinsically disordered proteins: what can we learn from kinetics? *Curr. Opin. Struct. Biol.* **36**, 18–24 (2016).

71. Uversky, V. N. Unusual biophysics of intrinsically disordered proteins. *Biochim. Biophys. Acta - Proteins Proteomics* **1834**, 932–951 (2013).

72. Uversky, V. N. A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *J. Biomol. Struct. Dyn.* **21**, 211–234 (2003).

73. Liu, Z. & Huang, Y. Advantages of proteins being disordered. *Protein Sci.* **23**, 539–550 (2014).

74. He, B. *et al.* Predicting intrinsic disorder in proteins: An overview. *Cell Res.* **19**, 929–949 (2009).

75. Worbs, M., Huber, R. & Wahl, M. C. Crystal structure of ribosomal protein L4 shows RNA-binding sites for ribosome incorporation and feedback control of the S10 operon. *EMBO J.* **19**, 807–818 (2000).

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

77. Jane Dyson, H. & Wright, P. E. [12] - Nuclear Magnetic Resonance Methods for Elucidation of Structure and Dynamics in Disordered States. in *Nuclear Magnetic Resonance of Biological Macromolecules - Part B* (eds. James, T. L., Dötsch, V. & Schmitz, U.) vol. 339 258–270 (Academic Press, 2001).

TH-3004_146106002

78. Alexandrescu, A. T., Abeygunawardana, C. & Shortle, D. Structure and Dynamics of a Denatured 131-Residue Fragment of Staphylococcal Nuclease: A Heteronuclear NMR Study. *Biochemistry* **33**, 1063–1072 (1994).

79. Habchi, J., Tompa, P., Longhi, S. & Uversky, V. N. Introducing protein intrinsic disorder. *Chem. Rev.* **114**, 6561–6588 (2014).

80. Showalter, S. A. Intrinsically Disordered Proteins: Methods for Structure and Dynamics Studies. in *eMagRes* 181–190 (John Wiley & Sons, Ltd, 2014).

81. Adler, A. J., Greenfield, N. J. & Fasman, G. D. B. T.-M. in E. [27] Circular dichroism and optical rotatory dispersion of proteins and polypeptides. in *Part D: Enzyme Structure* vol. 27 675–735 (Academic Press, 1973).

82. Kelly, S. M. & Price, N. C. The application of circular dichroism to studies of protein folding and unfolding. *Biochim. Biophys. Acta* **1338**, 161–185 (1997).

83. Jeganathan, S., von Bergen, M., Brutlach, H., Steinhoff, H.-J. & Mandelkow, E. Global hairpin folding of tau in solution. *Biochemistry* **45**, 2283–2293 (2006).

84. Tompa, P. & Kalmar, L. Power Law Distribution Defines Structural Disorder as a Structural Element Directly Linked with Function. *J. Mol. Biol.* **403**, 346–350 (2010).

85. Syme, C. D. *et al.* A Raman optical activity study of rheomorphism in caseins, synucleins and tau. *Eur. J. Biochem.* **269**, 148–156 (2002).

86. Ptitsyn, O. B. Structures of folding intermediates. *Curr. Opin. Struct. Biol.* **5**, 74–78 (1995).

87. Uversky, V. N. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry* **32**, 13288–13298 (1993).

88. Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z. & Dunker, A. K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **323**, 573–584 (2002).

TH-3004_146106002

89. Szollosi, E. *et al.* Intrinsic structural disorder of DF31, a drosophila protein of chromatin decondensation and remodeling activities. *J. Proteome Res.* **7**, 2291–2299 (2008).

90. Xie, H. *et al.* Functional Anthology of Intrinsic Disorder. 1. Biological Processes and Functions of Proteins with Long Disordered Regions. *J. Proteome Res.* **6**, 1882–1898 (2007).

91. Pietrosemoli, N., García-Martín, J. A., Solano, R. & Pazos, F. Genome-Wide Analysis of Protein Disorder in Arabidopsis thaliana: Implications for Plant Environmental Adaptation. *PLoS One* **8**, (2013).

92. Sun, X., Rikkerink, E. H. A., Jones, W. T. & Uversky, V. N. Multifarious roles of intrinsic disorder in proteins illustrate its broad impact on plant biology. *Plant Cell* **25**, 38–55 (2013).

93. Eriksson, S., Eremina, N., Barth, A., Danielsson, J. & Harryson, P. Membrane-Induced Folding of the Plant Stress Dehydrin Lti30 . *Plant Physiol.* **171**, 932–943 (2016).

94. Eriksson, S. K., Kutzer, M., Procek, J., Gröbner, G. & Harryson, P. Tunable membrane binding of the intrinsically disordered dehydrin Lti30, a cold-induced plant stress protein. *Plant Cell* **23**, 2391–2404 (2011).

95. Kovacs, D., Szabo, B., Pancsa, R. & Tompa, P. Intrinsically disordered proteins undergo and assist folding transitions in the proteome. *Arch. Biochem. Biophys.* **531**, 80–89 (2013).

96. Elshire, R. J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, 1–10 (2011).

97. Chakrabortee, S. *et al.* Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18073–18078 (2007).

98. Hundertmark, M. & Hincha, D. K. LEA (Late Embryogenesis Abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC Genomics* **9**, 1–22 (2008).

TH-3004_146106002

99. Kovacs, D. *et al.* Janus chaperones: assistance of both RNA- and protein-folding by ribosomal proteins. *FEBS Lett.* **583**, 88–92 (2009).

100. Ismail, A. M., Hall, A. E. & Close, T. J. Purification and partial characterization of a dehydrin involved in chilling tolerance during seedling emergence of cowpea. *Plant Physiol.* **120**, 237–244 (1999).

101. Hara, M., Terashima, S. & Kuboi, T. Characterization and cryoprotective activity of cold-responsive dehydrin from Citrus unshiu. *J. Plant Physiol.* **158**, 1333–1339 (2001).

102. Xie, C. *et al.* Overexpression of MtCAS31 enhances drought tolerance in transgenic Arabidopsis by reducing stomatal density. *New Phytol.* **195**, 124–135 (2012).

103. Close, T. J. Dehydrins: Emergence of a biochemical role of a family of plant dehydration proteins. *Physiol. Plant.* **97**, 795–803 (1996).

104. Malik, A. A., Veltri, M., Boddington, K. F., Singh, K. K. & Graether, S. P. Genome Analysis of Conserved Dehydrin Motifs in Vascular Plants. *Front. Plant Sci.* **8**, (2017).

105. Campbell, S. A. & Close, T. J. Dehydrins: genes, proteins, and associations with phenotypic traits. *New Phytol.* **137**, 61–74 (1997).

106. Dure III, L. A repeating 11-mer amino acid motif and plant desiccation. *Plant J.* **3**, 363–369 (1993).

107. Richard Strimbeck, G. Hiding in plain sight: the F segment and other conserved features of seed plant SK(n) dehydrins. *Planta* **245**, 1061–1066 (2017).

108. Koag, M.-C. *et al.* The K-segment of maize DHN1 mediates binding to anionic phospholipid vesicles and concomitant structural changes. *Plant Physiol.* **150**, 1503–1514 (2009).

109. Drira, M. *et al.* The K-Segments of the Wheat Dehydrin DHN-5 are Essential for the Protection of Lactate Dehydrogenase and β-Glucosidase Activities In Vitro. *Mol. Biotechnol.* **54**, 643–650 (2013).

TH-3004_146106002

110. Eriksson, S. K., Kutzer, M., Procek, J., Gröbner, G. & Harryson, P. Tunable Membrane Binding of the Intrinsically Disordered Dehydrin Lti30, a Cold-Induced Plant Stress Protein. *Plant Cell* **23**, 2391–2404 (2011).

111. Hughes, S. L. *et al.* The importance of size and disorder in the cryoprotective effects of dehydrins. *Plant Physiol.* **163**, 1376–1386 (2013).

112. Jiang, X. & Wang, Y. β-Elimination coupled with tandem mass spectrometry for the identification of in vivo and in vitro phosphorylation sites in maize dehydrin DHN1 protein. *Biochemistry* **43**, 15567–15576 (2004).

113. Goday, A. *et al.* The maize abscisic acid-responsive protein Rab17 is located in the nucleus and interacts with nuclear localization signals. *Plant Cell* **6**, 351–360 (1994).

114. Graether, S. P. & Boddington, K. F. Disorder and function: A review of the dehydrin protein family. *Front. Plant Sci.* **5**, 1–12 (2014).

115. CHATELAIN, E. *et al.* Temporal profiling of the heat-stable proteome during late maturation of Medicago truncatula seeds identifies a restricted subset of late embryogenesis abundant proteins associated with longevity. *Plant, Cell \& Environ.* **35**, 1440–1455 (2012).

116. Agarwal, T. *et al.* Different dehydrins perform separate functions in Physcomitrella patens. *Planta* **245**, 101–118 (2017).

117. Hara, M. *et al.* The Arabidopsis KS-type dehydrin recovers lactate dehydrogenase activity inhibited by copper with the contribution of His residues. *Plant Sci.* **245**, 135–142 (2016).

118. Zaccai, N. R., Serdyuk, I. N. & Zaccai, J. *Methods in molecular biophysics: structure, dynamics, function for biology and medicine*. (Cambridge University Press, 2017).

119. Hunt, H. D. & Simpson, W. T. Spectra of Simple Amides in the Vacuum Ultraviolet1. *J. Am. Chem. Soc.* **75**, 4540–4543 (1953).

120. Ham, J. S. & Platt, J. R. Far U.V. spectra of peptides. *J. Chem. Phys.* **20**, 335–336

TH-3004_146106002

(1952).

121. Imahori, K. & Tanaka, J. Ultraviolet absorption spectra of poly-L-glutamic acid. *J. Mol. Biol.* **1**, 359–364 (1959).

122. Bent, D. V & Hayon, E. Excited state chemistry of aromatic amino acids and related peptides. I. Tyrosine. *J. Am. Chem. Soc.* **97**, 2599–2606 (1975).

123. Creed, D. THE PHOTOPHYSICS AND PHOTOCHEMISTRY OF THE NEAR-UV ABSORBING AMINO ACIDS–I. TRYPTOPHAN AND ITS SIMPLE DERIVATIVES. *Photochem. Photobiol.* **39**, 537–562 (1984).

124. Longworth, J. W. Luminescence of Polypeptides and Proteins. *Excit. States Proteins Nucleic Acids* 319–484 (1971).

125. Antosiewicz, J. M. & Shugar, D. UV-Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 1: basic principles and properties of tyrosine chromophore. *Biophys. Rev.* **8**, 151–161 (2016).

126. Wetlaufer, D. B. Ultraviolet spectra Of Proteins and Amino Acids. in (eds. Anfinsen, C. B., Bailey, K., Anson, M. L. & Edsall, J. T. B. T.-A. in P. C.) vol. 17 303–390 (Academic Press, 1963).

127. OTEY, M. C. & GREENSTEIN, J. P. Studies on polycysteine peptides and proteins. II. Apparent dissociation constants, and ultra-violet and infrared absorption spectra of isomeric cystinylcystine peptides. *Arch. Biochem. Biophys.* **53**, 501–513 (1954).

128. Gutmann, F. Charge transfer complexes in biological systems. in (1997).

129. King, B. A., de Winter, A., McAnaney, T. B. & Boxer, S. G. Excited State Energy Transfer Pathways in Photosynthetic Reaction Centers. 4. Asymmetric Energy Transfer in the Heterodimer Mutant. *J. Phys. Chem. B* **105**, 1856–1862 (2001).

130. Balzani, V., Bergamini, G., Campagna, S. & Puntoriero, F. Photochemistry and photophysics of coordination compounds: overview and general concepts. *Photochem. photophysics Coord. Compd. I* 1–36 (2007).

131. Aliverti, A., Curti, B. & Vanoni, M. A. Identifying and quantitating FAD and FMN

TH-3004_146106002

in simple and in iron-sulfur-containing flavoproteins. *Methods Mol. Biol.* **131**, 9–23 (1999).

132. Arnstein, H. R. V. Data for biochemical research (Third Edition) : by R.M.C. Dawson, D.C. Elliott, W.H. Elliott and K.M. Jones Oxford University Press; Oxford, 1986 xii + 580 pages. £35.00. *FEBS Lett.* **234**, 506 (1988).

133. Prütz, W. A., Land, E. J. & Sloper, R. W. Charge transfer in peptides. Effects of temperature, peptide length and solvent conditions upon intramolecular one-electron reactions involving tryptophan and tyrosine. *J. Chem. Soc. Faraday Trans. 1 Phys. Chem. Condens. Phases* **77**, 281–292 (1981).

134. Arikuma, Y., Nakayama, H., Morita, T. & Kimura, S. Electron hopping over 100 Å along an α helix. *Angew. Chemie Int. Ed.* **49**, 1800–1804 (2010).

135. Ryzhkina, I. S. *et al.* Self-organization and properties of dispersed systems based on dilute aqueous solutions of (S) - and (R) -lysine. *Ital. Oral Surg.* **28**, 66–69 (2018).

136. Homchaudhuri, L. & Swaminathan, R. Novel Absorption and Fluorescence Characteristics of L-Lysine. *Chem. Lett.* **30**, 844–845 (2001).

137. Homchaudhuri, L. & Swaminathan, R. Near Ultraviolet Absorption Arising from Lysine Residues in Close Proximity: A Probe to Monitor Protein Unfolding and Aggregation in Lysine-Rich Proteins. *Bull. Chem. Soc. Jpn.* **77**, 765–769 (2004).

138. Chen, X. *et al.* Prevalent intrinsic emission from nonaromatic amino acids and poly(amino acids). *Sci. China Chem.* **61**, 351–359 (2018).

139. Chai, B., Zheng, J., Zhao, Q. & Pollack, G. H. Spectroscopic Studies of Solutes in Aqueous Solution. *J. Phys. Chem. A* **112**, 2242–2247 (2008).

140. Guptasarma, P. Solution-state characteristics of the ultraviolet A-induced visible fluorescence from proteins. *Arch. Biochem. Biophys.* **478**, 127–129 (2008).

141. Tikhonova, T. N. *et al.* Dissection of the deep-blue autofluorescence changes accompanying amyloid fibrillation. *Arch. Biochem. Biophys.* **651**, 13–20 (2018).

142. Shukla, A. *et al.* A novel UV laser-induced visible blue radiation from protein

TH-3004_146106002

crystals and aggregates: scattering artifacts or fluorescence transitions of peptide electrons delocalized through hydrogen bonding? *Arch. Biochem. Biophys.* **428**, 144–153 (2004).

143. Villa, A. M., Doglia, S. M., De Gioia, L., Bertini, L. & Natalello, A. Anomalous Intrinsic Fluorescence of HCl and NaOH Aqueous Solutions. *J. Phys. Chem. Lett.* **10**, 7230–7236 (2019).

144. Prasad, S. *et al.* Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. *Chem. Sci.* **8**, 5416–5433 (2017).

145. Mandal, I., Paul, S. & Venkatramani, R. Optical backbone-sidechain charge transfer transitions in proteins sensitive to secondary structure and modifications. *Faraday Discuss.* **207**, 115–135 (2018).

146. Kumar, A., Ahari, D., Priyadarshi, A., Ziauddin Ansari, M. & Swaminathan, R. Weak Intrinsic Luminescence in Monomeric Proteins Arising from Charge Recombination. *J. Phys. Chem. B* **124**, 2731–2746 (2020).

147. Bhattacharya, A., Prajapati, R., Chatterjee, S. & Mukherjee, T. K. Concentration-Dependent Reversible Self-Oligomerization of Serum Albumins through Intermolecular β-Sheet Formation. *Langmuir* **30**, 14894–14904 (2014).

148. Bhattacharya, A. *et al.* Direct Evidence of Intrinsic Blue Fluorescence from Oligomeric Interfaces of Human Serum Albumin. *Langmuir* **33**, 10606–10615 (2017).

149. Jepson, S. G. & Close, T. J. Purification of a maize dehydrin protein expressed in escherichia coli. *Protein Expression and Purification* vol. 6 632–636 (1995).

150. Edelheit, O., Hanukoglu, A. & Hanukoglu, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* **9**, 1–8 (2009).

151. Nelson, M. & McClelland, M. Use of DNA methyltransferase/endonuclease enzyme combinations for megabase mapping of chromosomes. *Methods Enzymol.* **216**, 279–303 (1992).

TH-3004_146106002

152. Braman, J., Papworth, C. & Greener, A. Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Mol. Biol.* **57**, 31–44 (1996).

153. Cid, M.-M. & Bravo, J. *Structure Elucidation in Organic Chemistry: The Search for the Right Tools*. (John Wiley & Sons, 2014).

154. Hof, M. Basics of optical spectroscopy. *Handb. Spectrosc.* **1**, 39–47 (2003).

155. Cantor, C. R. & Schimmel, P. R. Techniques for the study of biological structure and function. (1980).

156. Lakowicz, J. . *Principles of Fluorescence Spectroscopy*. (Springer , Boston, MA, 2006).

157. Walla, P. J. *Modern biophysical chemistry: detection and analysis of biomolecules*. (John Wiley & Sons, 2014).

158. Valeur, B. & Berberan-Santos, M. N. *Molecular fluorescence: principles and applications*. (John Wiley & Sons, 2012).

159. Vacic, V., Uversky, V. N., Dunker, A. K. & Lonardi, S. Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *BMC Bioinformatics* **8**, 211 (2007).

160. Kozlowski, L. P. & Bujnicki, J. M. MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. *BMC Bioinformatics* **13**, 111 (2012).

161. Callebaut, I. *et al.* Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cell. Mol. Life Sci.* **53**, 621–645 (1997).

162. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual.* (Cold spring harbor laboratory press, 1989).

163. LOWRY, O. H., ROSEBROUGH, N. J., FARR, A. L. & RANDALL, R. J. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265–275 (1951).

164. Wiederschain, G. Y. The Molecular Probes handbook. A guide to fluorescent probes and labeling technologies. *Biochem.* **76**, 1276–1277 (2011).

TH-3004_146106002

165. Hudson, E. N. & Weber, G. Synthesis and characterization of two fluorescent sulfhydryl reagents. *Biochemistry* **12**, 4154–4161 (1973).

166. Brouwer, A. M. Standards for photoluminescence quantum yield measurements in solution (IUPAC Technical Report). *Pure Appl. Chem.* **83**, 2213–2228 (2011).

167. Swaminathan, R., Periasamy, N., Udgaonkar, J. B. & Krishnamoorthy, G. Molten Globule-like Conformation of Barstar: A Study by Fluorescence Dynamics. *J. Phys. Chem.* **98**, 9270–9278 (1994).

168. Swaminathan, R., Krishnamoorthy, G. & Periasamy, N. Similarity of fluorescence lifetime distributions for single tryptophan proteins in the random coil state. *Biophys. J.* **67**, 2013–2023 (1994).

169. Swaminathan, R. & Periasamy, N. Analysis of fluorescence decay by the maximum entropy method: Influence of noise and analysis parameters on the width of the distribution of lifetimes. *Proc. Indian Acad. Sci. - Chem. Sci.* **108**, 39–49 (1996).

170. Wu, P. & Brand, L. Resonance energy transfer: methods and applications. *Anal. Biochem.* **218**, 1–13 (1994).

171. Lobley, A., Whitmore, L. & Wallace, B. A. DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* **18**, 211–212 (2002).

172. Whitmore, L. & Wallace, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668–W673 (2004).

173. Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **89**, 392–400 (2008).

174. Sreerama, N. & Woody, R. W. Structural composition of βI- and βII-proteins. *Protein science : a publication of the Protein Society* vol. 12 384–388 (2003).

175. Cuevas-Velazquez, C. L., Saab-Rincón, G., Reyes, J. L. & Covarrubias, A. A. The

TH-3004_146106002

unstructured n-terminal region of Arabidopsis group 4 late embryogenesis abundant (LEA) proteins is required for folding and for chaperone-like activity under water deficit. *J. Biol. Chem.* **291**, 10893–10903 (2016).

176. Drira, M. *et al.* The K-segments of the wheat dehydrin DHN-5 are essential for the protection of lactate dehydrogenase and β-glucosidase activities in vitro. *Mol. Biotechnol.* **54**, 643–650 (2013).

177. Goyal, K., Walton, L. J. & Tunnacliffe, A. LEA proteins prevent protein aggregation due to water stress. *Biochem. J.* **388**, 151–157 (2005).

178. Antosiewicz, J. M. & Shugar, D. UV-Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 2: selected applications. *Biophys. Rev.* **8**, 163–177 (2016).

179. Ghisaidoobe, A. B. T. & Chung, S. J. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: A focus on förster resonance energy transfer techniques. *Int. J. Mol. Sci.* **15**, 22518–22538 (2014).

180. Rami, B. R., Krishnamoorthy, G. & Udgaonkar, J. B. Dynamics of the core tryptophan during the formation of a productive molten globule intermediate of barstar. *Biochemistry* **42**, 7986–8000 (2003).

181. Chamberlain, C. & Hahn, K. M. Watching proteins in the wild: Fluorescence methods to study protein dynamics in living cells. *Traffic* **1**, 755–762 (2000).

182. Komatsu, T. *et al.* Real-Time Measurements of Protein Dynamics Using Fluorescence Activation-Coupled Protein Labeling Method. *J. Am. Chem. Soc.* **133**, 6745–6751 (2011).

183. TEALE, F. W. & WEBER, G. Ultraviolet fluorescence of the aromatic amino acids. *Biochem. J.* **65**, 476–482 (1957).

184. Guy, J. *et al.* Convergent Preparation and Photophysical Characterization of Dimaleimide Dansyl Fluorogens: Elucidation of the Maleimide Fluorescence Quenching Mechanism. *J. Am. Chem. Soc.* **129**, 11969–11977 (2007).

TH-3004_146106002

185. Kumar, A. *et al.* Role of Charged Amino Acids in Sullying the Fluorescence of Tryptophan or Conjugated Dansyl Probe in Monomeric Proteins. *Biochemistry* **61**, 339–353 (2022).

186. Ansari, M. Z. & Swaminathan, R. Structure and dynamics at N- and C-terminal regions of intrinsically disordered human c-Myc PEST degron reveal a pH-induced transition. *Proteins Struct. Funct. Bioinforma.* **88**, 889–909 (2020).

187. Clarke, M. W. *et al.* Structural and functional insights into the cryoprotection of membranes by the intrinsically disordered dehydrins. *J. Biol. Chem.* **290**, 26900–26913 (2015).

188. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. (Kluwer Acadmic, 1999).

189. Gregoret, L. M. & Sauer, R. T. Tolerance of a protein helix to multiple alanine and valine substitutions. *Fold. Des.* **3**, 119–126 (1998).

190. Nagano, N., Ota, M. & Nishikawa, K. Strong hydrophobic nature of cysteine residues in proteins. *FEBS Lett.* **458**, 69–71 (1999).

191. Beechem, J. M. & Brand, L. TIME-RESOLVED FLUORESCENCE OF PROTEINS. *Annu. Rev. Biochem.* **54**, 43–71 (1985).

192. Brand, L. & Gohlke, J. R. Fluorescence Probes for Structure. *Annu. Rev. Biochem.* **41**, 843–868 (1972).

193. Zhdanova, N. G. *et al.* Tyrosine fluorescence probing of the surfactant-induced conformational changes of albumin. *Photochem. Photobiol. Sci.* **14**, 897–908 (2015).

194. Vivian, J. T. & Callis, P. R. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.* **80**, 2093–2109 (2001).

195. Chen, Y. & Barkley, M. D. Toward understanding tryptophan fluorescence in proteins. *Biochemistry* **37**, 9976–9982 (1998).

196. Steiner, R. F. Fluorescence Anisotropy: Theory and Applications. in *Topics in Fluorescence Spectroscopy: Principles* (ed. Lakowicz, J. R.) 1–52 (Springer US, 2002).

TH-3004_146106002

197. Demchenko, A. P. *Ultraviolet spectroscopy of proteins*. (Springer Science & Business Media, 2013).

198. Mateos-Gil, P., Tsortos, A., Vélez, M. & Gizeli, E. Monitoring structural changes in intrinsically disordered proteins using QCM-D: application to the bacterial cell division protein ZipA. *Chem. Commun.* **52**, 6541–6544 (2016).

199. Ansari, M. Z. *et al.* Protein charge transfer absorption spectra: an intrinsic probe to monitor structural and oligomeric transitions in proteins. *Faraday Discuss.* **207**, 91–113 (2018).

200. Fasman, G. Fasman, Practical handbook of biochemistry and molecular biology, CRC Press , New York. (1992).

201. Howell, J.R Menguc, M Pinar, Siegel, R. *Thermal radiation heat transfer*. *CRC* (CRC, 2002).

202. H.C. van de Hulst. Light scattering by small particles. By H. C. van de Hulst. New York (John Wiley and Sons), London (Chapman and Hall), 1957. Pp. xiii, 470; 103 Figs.; 46 Tables. 96s. *Q. J. R. Meteorol. Soc.* **84**, 198–199 (1958).

203. Rayleigh, Lord. XXXIV. On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky. *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **47**, 375–384 (1899).

204. Han, X. X., Huang, G. G., Zhao, B. & Ozaki, Y. Label-Free Highly Sensitive Detection of Proteins in Aqueous Solutions Using Surface-Enhanced Raman Scattering. *Anal. Chem.* **81**, 3329–3333 (2009).

205. Anastas, P. & Eghbali, N. Green Chemistry: Principles and Practice. *Chem. Soc. Rev.* **39**, 301–312 (2010).

206. Kasha, M. Characterization of electronic transitions in complex molecules. *Discuss. Faraday Soc.* **9**, 14–19 (1950).

207. Demchenko, A. P., Tomin, V. I. & Chou, P.-T. Breaking the Kasha Rule for More Efficient Photochemistry. *Chem. Rev.* **117**, 13353–13381 (2017).

TH-3004_146106002

208. Scuppa, S., Orian, L., Donoli, A., Santi, S. & Meneghetti, M. Anti-Kasha's Rule Fluorescence Emission in (2-Ferrocenyl)indene Generated by a Twisted Intramolecular Charge-Transfer (TICT) Process. *J. Phys. Chem. A* **115**, 8344–8349 (2011).

209. Shen, X.-B. *et al.* Excitation-wavelength-dependent photoluminescence of silicon nanoparticles enabled by adjustment of surface ligands. *Chem. Commun.* **54**, 4947–4950 (2018).

210. Eftink, M. R. Fluorescence techniques for studying protein structure. *Methods Biochem. Anal.* **35**, 127–205 (1991).

211. Eftink, M. R., Jia, J., Hu, D. & Ghiron, C. A. Fluorescence Studies with Tryptophan Analogs: Excited State Interactions Involving the Side Chain Amino Group. *J. Phys. Chem.* **99**, 5713–5723 (1995).

212. Chen, J. Intrinsically disordered p53 extreme C-terminus binds to S100B(ββ) through 'fly-casting'. *J. Am. Chem. Soc.* **131**, 2088–2089 (2009).

213. Dosnon, M. *et al.* Demonstration of a Folding after Binding Mechanism in the Recognition between the Measles Virus NTAIL and X Domains. *ACS Chem. Biol.* **10**, 795–802 (2015).

214. Uversky, V. N. Functional roles of transiently and intrinsically disordered regions within proteins. *FEBS J.* **282**, 1182–1189 (2015).

215. Marín, M. & Ott, T. Intrinsic disorder in plant proteins and phytopathogenic bacterial effectors. *Chem. Rev.* **114**, 6912–6932 (2014).

216. Tunnacliffe, A. & Wise, M. J. The continuing conundrum of the LEA proteins. *Naturwissenschaften* **94**, 791–812 (2007).

217. Kalthoff, C. A novel strategy for the purification of recombinantly expressed unstructured protein domains. *J. Chromatogr. B* **786**, 247–254 (2003).

218. Livernois, A. M., Hnatchuk, D. J., Findlater, E. E. & Graether, S. P. Obtaining highly purified intrinsically disordered protein by boiling lysis and single step ion exchange.

TH-3004_146106002

Anal. Biochem. **392**, 70–76 (2009).

219. KrishnaKumar, V. G. & Gupta, S. Simplified method to obtain enhanced expression of tau protein from E. coli and one-step purification by direct boiling. *Prep. Biochem. Biotechnol.* **47**, 530–538 (2017).

220. Hanin, M. *et al.* Plant dehydrins and stress tolerance: Versatile proteins for complex mechanisms. *Plant Signal. Behav.* **6**, 1503–1509 (2011).

221. Rahman, L. N. *et al.* Interactions of intrinsically disordered Thellungiella salsuginea dehydrins TsDHN-1 and TsDHN-2 with membranes - Synergistic effects of lipid composition and temperature on secondary structure. *Biochem. Cell Biol.* **88**, 791–807 (2010).

222. Tolleter, D. *et al.* Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* **19**, 1580–1589 (2007).

223. Liu, Y., Song, Q., Li, D., Yang, X. & Li, D. Multifunctional Roles of Plant Dehydrins in Response to Environmental Stresses. *Front. Plant Sci.* **8**, 1018 (2017).

224. Koag, M. C. *et al.* The K-segment of maize DHN1 mediates binding to anionic phospholipid vesicles and concomitant structural changes. *Plant Physiol.* **150**, 1503–1514 (2009).

225. Zhang, M., Pascal, J. M. & Zhang, J. F. Unstructured to structured transition of an intrinsically disordered protein peptide in coupling Ca2+-sensing and SK channel activation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4828–4833 (2013).

226. Buck, M. Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* **31**, 297–355 (1998).

227. Luo, P. & Baldwin, R. L. Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36**, 8413–8421 (1997).

228. Uversky, V. N. Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, Counter ions, membranes, binding partners,

TH-3004_146106002

osmolytes, and macromolecular crowding. *Protein J.* **28**, 305–325 (2009).

229. Jemth, P. *et al.* Structure and dynamics conspire in the evolution of affinity between intrinsically disordered proteins. *Sci. Adv.* **4**, (2018).

230. Ferreon, A. C. M., Ferreon, J. C., Wright, P. E. & Deniz, A. A. Modulation of allostery by protein intrinsic disorder. *Nature* **498**, 390–394 (2013).

231. Koag, M. C., Fenton, R. D., Wilkens, S. & Close, T. J. The binding of maize DHN1 to lipid vesicles. Gain of structure and lipid specificity. *Plant Physiol.* **131**, 309–316 (2003).

232. Hughes, S. & Graether, S. P. Cryoprotective mechanism of a small intrinsically disordered dehydrin protein. *Protein Sci.* **20**, 42–50 (2011).

233. Díaz-Villanueva, J. F., Díaz-Molina, R. & García-González, V. Protein folding and mechanisms of proteostasis. *Int. J. Mol. Sci.* **16**, 17193–17230 (2015).

234. Dogan, J., Jonasson, J., Andersson, E. & Jemth, P. Binding Rate Constants Reveal Distinct Features of Disordered Protein Domains. *Biochemistry* **54**, 4741–4750 (2015).

235. Kovacs, D., Szabo, B., Pancsa, R. & Tompa, P. Intrinsically disordered proteins undergo and assist folding transitions in the proteome. *Arch. Biochem. Biophys.* **531**, 80–89 (2013).

236. Lawrence, C. W., Kumar, S., Noid, W. G. & Showalter, S. A. Role of ordered proteins in the folding-upon-binding of intrinsically disordered proteins. *J. Phys. Chem. Lett.* **5**, 833–838 (2014).

237. Dosnon, M. *et al.* Demonstration of a folding after binding mechanism in the recognition between the measles virus NTAIL and X domains. *ACS Chem. Biol.* **10**, 795–802 (2015).

238. Keppel, T. R., Howard, B. A. & Weis, D. D. Mapping unstructured regions and synergistic folding in intrinsically disordered proteins with amide H/D exchange mass spectrometry. *Biochemistry* **50**, 8722–8732 (2011).

TH-3004_146106002

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

240. Stryer, L. Fluorescence spectroscopy of proteins. *Science (80-.).* **162**, 526–533 (1968).

241. Joder, K., Kiefhaber, T. & Mo, A. Correction for Moglich et al., End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. *Proc. Natl. Acad. Sci.* **105**, 6787–6787 (2008).

242. Boyer, S. B. & Slesinger, P. A. Probing novel GPCR interactions using a combination of FRET and TIRF. *Commun. Integr. Biol.* **3**, 343–346 (2010).

243. Alkudaisi, N., Russell, B. A., Jachimska, B., Birch, D. J. S. & Chen, Y. Detecting lysozyme unfolding: Via the fluorescence of lysozyme encapsulated gold nanoclusters. *J. Mater. Chem. B* **7**, 1167–1175 (2019).

244. Cuevas-Velazquez, C. L., Rendón-Luna, D. F. & Covarrubias, A. A. Dissecting the cryoprotection mechanisms for dehydrins. *Front. Plant Sci.* **5**, 1–6 (2014).

245. Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta* **1804**, 1231–1264 (2010).

246. Gibbs, E. B. & Showalter, S. A. Quantitative biophysical characterization of intrinsically disordered proteins. *Biochemistry* **54**, 1314–1326 (2015).

247. Sickmeier, M. *et al.* DisProt: the Database of Disordered Proteins. *Nucleic Acids Res.* **35**, D786–D793 (2007).

248. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

249. Uversky, V. N. Intrinsically disordered proteins and their 'Mysterious' (meta)physics. *Front. Phys.* **7**, 8–23 (2019).

250. Van Der Lee, R. *et al.* Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 (2014).

TH-3004_146106002

251. Yruela, I. & Contreras-Moreira, B. Protein disorder in plants: a view from the chloroplast. *BMC Plant Biol.* **12**, 165 (2012).

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी • Indian Institute of Technology Guwahati •

TH-3004_146106002