Monitoring Protein Unfolding and DNA Binding using Protein Charge Transfer Spectra

A thesis submitted in partial fulfilment of the
Requirements for the degree of

DOCTOR OF PHILOSOPHY

By

ANURAG PRIYADARSHI

Roll No. 156106026

Department of Biosciences and Bioengineering,

Indian Institute of Technology Guwahati,

Guwahati, Assam- 781039, India

August 2022

INDIAN INSTITUTE OF TECHNOLOGY

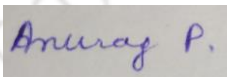
GUWAHATI, Assam, India

Department of Biosciences and Bioengineering

STATEMENT

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

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



IIT Guwahati — Anurag Priyadarshi

August, 2022

INDIAN INSTITUTE OF TECHNOLOGY

GUWAHATI, Assam, India

Department of Biosciences and Bioengineering

CERTIFICATE

It is certified that the work described in this thesis titled "***Monitoring Protein Unfolding and DNA Binding using Protein Charge Transfer Spectra"***, carried out by **Anurag Priyadarshi** for the award of the 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
Department of BSBE
IIT Guwahati

Dedicated to my

Family

Acknowledgements

I am grateful for the opportunities life has provided me so that I was able to pursue my doctorate degree at IIT Guwahati. I would like to express my sincere gratitude to the people who have been a part of my personal and professional life during this endeavor.

First and foremost, I am extremely grateful to my supervisor Prof. R. Swaminathan for giving me the opportunity to pursue my Ph.D. in his research group and for his constant faith in me. His moral support, motivation, and enthusiasm helped me immensely during tough times. His scientific insights related to my thesis work were extremely valuable. As a mentor he has brought out the best in me by encouraging me to pursue new research ideas and guiding me in the correct direction.

I am extremely thankful to my Doctoral Committee members Prof. Sandip Paul, Prof. B. Anand, Prof. Lalit Mohan Pandey, and past DC member Prof. Pranjal Chandra for their critical assessment of my Ph.D. work, valuable suggestions, and constant encouragement which helped me develop crucial critical thinking and problem solving skills required.

I am extremely thankful to Prof. Cecilia Tommos from Texas A&M University, USA, for providing the plasmid constructs of α_3C and α_3W. I am also thankful to Prof DN Rao, Indian Institute of Science, Bengaluru for allowing me to work in his lab and Dr. Sumedha Kondekar for helping me in the purification of HpTopoI protein.

Many experiments were conducted using the DCIF facilities at the Department of Biosciences and Bioengineering, IIT Guwahati and I am thankful for the cooperation of staff members at the department. I spent considerable amount of time at Central Instruments Facility, IIT Guwahati both as an operator of the 600 MHz NMR spectroscopy instrument and as a user of Mass spectrometry (MALDI) instrument. IIT Guwahati provided the facilities which allowed me to pursue my academic goals and the Ministry of Human Resource Development (MHRD), India, gave the financial support.

I am also thankful to the IIT Guwahati administration for prompt reopening of the research labs after the COVID lockdowns. Access to the research facilities was maintained even in difficult circumstances allowing students to perform experiments.

I am thankful to my lab seniors Tulsi Swain, Saumya Prasad, Shrutidhara Biswas, Mohd. Ziauddin Ansari, Amrendra Kumar, and Dileep Ahari for their help during my Ph.D. Sanjana, Ekramul, Alka, Heeramoni, Nuzelu, Nayanjyoti, Garima, Subhajit, Abheek, Harikrishna, Simangka,

Himanshi, Sahil, Kartik, Junaid, and Niteesh were very helpful labmates and I hopefully get a chance to work with them in the future. I am especially thankful to Amrendra Kumar for his help in learning the techniques used in my Ph.D and Ekramul Alom for help with computational calculations used in thesis work.

I am grateful to my friends Aman, Abhijeet, Jon, Avishek, Nayan, Sheel, and Nagendra who were an integral part of my stay at IIT Guwahati. Pratik Nag and Dr. Palash Bortamuly went out of their way to help me whenever I needed it. Mr Amit Thakur of Jaldhara Co. was always reliable when chemicals were required on an urgent basis.

My parents have given me unconditional love and support. It is their guidance which has helped me in becoming a better person and better at my job. I am thankful to my family members, especially my younger brother Piyush for being a constant source of encouragement and motivation. My friends from NISER, Bhubaneswar were always available for me when I needed them.

It is an impossible task to name all the people who have helped me throughout this academic journey. I am extremely thankful to each one of them for placing their trust in me.

Anurag Priyadarshi *August 2022*

LIST OF ABBREVIATIONS

AIE	Aggregation Induced Emission
ANS	8-anilino-1-napthalensulphonate
BSA	Bovine Serum Albumin
CD	Circular dichroism
ChIP	Chromatin Immunoprecipitation
CT	Charge transfer
CTE	Clustering Triggered Emission
Dansyl	5-(DimethylAmino)Naphthalene-1-SulfonYL
dbAF	deep-blue Autofluorescence
DHN1	Dehydrin 1
DLS	Dynamic Light Scattering
DNase I	Deoxyribonuclease I
DPA	9,10-diphenylanthracene
EMSA	Electrophoretic Mobility Shift Assay
EtBr	Ethidium Bromide
FAD	Flavin Adenine Dinucleotide
FRET	Förster resonance energy transfer
FTIR	Fourier-transform Infrared
FWHM	Full Width at Half Maximum
gDNA	genomic Deoxyribonucleic acid
gDNA	Calf thymus genomic DNA
GdnHCl	Guanidinium chloride
HEWL	Hen Egg White Lysozyme
HOMO	Highest Occupied Molecular Orbital
HSA	Human Serum Albumin
IAEDANS	5-({2-[(iodoacetyl)amino]ethyl}amino)naphthalene-1-sulfonic acid

IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
IRF	Instrument Response Function
ITC	Isothermal Titration Calorimetry
LUMO	Lowest Unoccupied Molecular Orbital
MD	Molecular Dynamics
MEM	Maximum Entropy Method
MRE	Mean residual ellipticity
NAD	Nicotinamide Adenine Dinucleotide
NATA	N-Acetyl-L-tryptophanamide
NLLS	Nonlinear Least Squares
NMR	Nuclear Magnetic Resonance
PBS-CT	Peptide Backbone to Sidechain Charge Transfer
PDB	Protein Data Bank
PET-FCS	Photoinduced Electron Transfer-Fluorescence Correlation Spectroscopy
PRM	Protamine
ProCharTS	Protein Charge Transfer Spectra
QY	Quantum yield
SAXS	Small Angle X-ray Scattering
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
SPR	Surface Plasmon Resonance
SS-CT	Sidechain to Sidechain Charge Transfer
TCEP	Tris(2-CarboxyEthyl)Phosphine
TCSPC	Time-Correlated Single Photon Counting
TDDFT	Time-Dependent Density-Functional Theory
ThT	Thioflavin T
UV-Vis	Ultraviolet-Visible

THESIS ABSTRACT

Molecular spectroscopy is widely used in the characterization of protein structure and function relationship.[1–3] Protein folding and DNA-protein binding are important processes involved in multiple biological phenomena.[4,5] The tool of UV-Visible spectroscopy is routinely used to characterize protein unfolding[6] and less frequently to monitor DNA-protein binding.[7] UV-Vis spectroscopy involves using intrinsic and extrinsic chromophores to study protein unfolding and DNA-protein binding. Aromatic amino acid residues are the intrinsic chromophores used in these studies. Proteins are not expected to show electronic absorption or emission in the near UV-Visible region in the absence of aromatic amino acid residues and cofactors.[8] However, evidence of novel protein absorbance and luminescence in the near UV-Visible region independent of the presence of aromatic amino acid residues has now been presented by multiple research groups.[9–11]

One of the earliest reports of absorbance from non-aromatic amino acids was made when concentrated solutions of L-Lysine.HCl was shown to absorb near UV wavelengths and emit blue luminescence by Swaminathan *et al* in 2001.[12] In a follow-up study, dilute solutions of poly-L-Lysine.HCl, calf thymus histone, and Human Serum Albumin (HSA) were reported to have novel absorbance and luminescence like observed from Lys solutions.[13] To systematically investigate the nature of the chromophore involved in the observed absorbance and luminescence, computational and UV-Vis spectroscopy studies were done on protein α_3C (rich in Lys and Glu amino acid residues but devoid of aromatic amino acid residues).[9] By employing Time-dependent density-functional theory (TDDFT) calculations, it was postulated that charge transfer involving charged atoms in the Lys/Glu sidechain and the peptide backbone gave rise to the novel absorbance observed in the 250—800 nm region for protein α_3C. The observed α_3C spectra was attributed to this previously unknown intrinsic chromophore, and the discovered phenomenon was termed Protein Charge Transfer Spectra (ProCharTS) by Prasad *et al*.[9]

Two types of photoinduced charge transfer were observed in the protein α_3C using TDDFT calculations by Prasad *et al*: (1) peptide backbone to sidechain charge transfer (PBS-CT) and (2) sidechain to sidechain charge transfer (SS-CT). PBS-CT involved charge transfer from the peptide backbone to Lys sidechain in Lys-Lys close contacts (3—4 Å distance). In Glu-Glu close contacts (3.5—5 Å distance), the direction of charge transfer was from Glu sidechain to the peptide backbone. SS-CT involved charge transfer from the carboxyl group in Glu sidechain to the amino group in Lys sidechain of a Glu-Lys pair at a moderate distance from each other (5—6 Å). In a subsequent study by Mandal *et al*, TDDFT calculations were used to show PBS-CT in other charged amino acid residues Asp, Arg, His, and the phosphorylated forms of Ser, Thr, and Tyr.[14] Importantly, the presence of PBS-CT for the amino acid residue Arg was also reported.

Weak luminescence from multiple charged monomeric proteins was reported recently by our group with characteristic features of the observed blue luminescence including low to moderate quantum yields, excitation-emission spectral overlap among the multiple proteins and charged

amino acids studied, large Stokes shifts, and similar mean lifetimes for luminescence intensity decays.[15]

The proteins α_3C, α_3W, and PRM are rich in charged amino acid residues and were hence chosen to be investigated for the presence of ProCharTS as the first objective of the thesis work. α_3W is the single Trp-containing variant of α_3C with Cys-34 replaced by Trp. The absorbance and luminescence of the proteins were measured, and the photophysical characteristics of novel luminescence such as Stokes shifts and the quantum yields were analysed. The quantum yields of α_3C and α_3W upon excitation at 355 nm were 0.0040 and 0.0037, whereas PRM quantum yield at 270 nm was 0.0108. The low quantum yields of the observed ProCharTS luminescence highlighted why only a few reports of the phenomena have been made till now. α_3C and α_3W emission had moderate Stokes shifts of 3090—6136 cm^{-1}, and PRM luminescence was characterized by high Stokes shifts (7609—16415 cm^{-1}).

Solvent parameters such as pH extremes and the presence of salt are expected to change the net charge and possibly the proximity of charged amino acid residues in proteins. The effects of solvent parameters on α_3C, α_3W, and PRM ProCharTS absorbance and luminescence were characterized. The correlation of changes in ProCharTS absorbance was done with structural transitions in α_3C and α_3W monitored by far UV CD spectroscopy. The effect of changes in solvent parameters on ProCharTS luminescence was less straightforward. An understanding of the effect of solvent parameters on ProCharTS was essential before investigating the potential applications of the novel intrinsic chromophore.

The contribution of protein charge transfer luminescence in the region of Trp emission was ascertained by our group recently[16] and was also investigated in the current study. In the presence of α_3C CT states, the fluorescence decay of NATA (Trp analog) was no longer a single exponential and fit to a sum of two exponentials. A short component was observed in α_3W Trp fluorescence decay upon analysis by the maximum entropy method, hinting toward the detection of CT luminescence along with Trp fluorescence.

The proteins HSA, α_3C, and α_3W are natively folded and rich in charged amino acid residues. Significant ProCharTS absorbance and luminescence from HSA has already been reported by Kumar *et al*.[15] α_3C ProCharTS absorbance in the near UV-Visible region was reported by Prasad et al,[9] and α_3W ProCharTS absorbance was observed in the present study. ProCharTS absorbance and luminescence depend on the number of charged amino acid sidechains in close contact. The number and nature of these contacts are expected to decrease and change substantially in the process of protein unfolding. Hence an attempt was made to use ProCharTS absorbance and luminescence to monitor protein unfolding.

The chemical denaturant-induced unfolding of α_3C, α_3W, and HSA was characterized using ProCharTS along with other conventional methods such as CD spectroscopy, α_3W and HSA Trp fluorescence, α_3C-Dansyl and HSA-Dansyl fluorescence. A decrease in α_3C, α_3W, and HSA ProCharTS absorbance was observed upon protein unfolding. ProCharTS intensity depends on the

number of charged amino acid residues in close proximity, and differences compared to the characterization of protein unfolding by the conventional methods could arise. Hence denaturation midpoint parameters obtained after analyzing protein unfolding by ProCharTS was compared to the conventional methods. Based on the quantitative analysis of GdnHCl-induced protein unfolding, it was deduced that the disruption of charged amino acid contacts preceded the loss of secondary structure and Trp/Dansyl solvent exposure. Additionally, the previously reported molten globule-like state in the presence of sub-denaturing GdnHCl concentrations was detected using HSA ProCharTS absorbance.[6]

Protamines (PRM) are sperm nuclear basic proteins that replace histone as the DNA packaging material in maturing spermatids. Protamines are abundant in basic amino acids (50—70 %), contain Arg clusters, and little to no negatively charged amino acid residues.[17] Salmon protamine (PRM) is abundant in arginines containing 21 Arg out of a total 32 amino acid residues. The abundance of Arg made PRM a suitable protein for the presence of PBS-CT (involving the Arg sidechain and the peptide backbone), leading to absorbance in the near UV region even in the absence of aromatic amino acid residues. The potential application of PRM ProCharTS absorbance to monitor genomic DNA-PRM binding was illustrated.

The condensation of genomic DNA (gDNA) by PRM has been illustrated earlier by multiple techniques including EMSA,[18] dynamic light scattering,[19] circular dichroism,[7] and UV-Visible spectroscopy.[7] PRM remains unstructured in solution, but clarity on PRM structure in nucleoprotamine complex is still elusive. Since ProCharTS is sensitive to the number of charged Arg sidechains in close contact, it was investigated whether PRM ProCharTS could be used to detect possible structure gain upon DNA binding.

PRM is a known DNA condensing agent and HEWL-DNA binding has also been reported earlier. Mixing calf thymus genomic DNA with PRM results in DNA condensation and nucleoprotein complex formation.[20] At low HEWL/PRM concentration, the nucleoprotein complex remained soluble, but in the presence of moderate or high protein concentration, precipitation of the nucleoprotein complex was observed. The decrease in protein/DNA concentration of gDNA-protein mixture supernatants was employed to monitor DNA-protein binding. Specifically, the decrease in PRM ProCharTS absorbance, HEWL absorbance, HEWL Trp fluorescence, and gDNA absorbance due to nucleoprotein complex precipitation was used to monitor DNA-protein binding. The precipitation of nucleoprotein complex upon gDNA binding with HEWL/PRM was also confirmed using agarose gel electrophoresis of DNA-protein mixture supernatants. The previously reported inability of BSA to condense DNA was also observed.[20] This study is thus the first to use ProCharTS to monitor DNA-protein binding.

To summarize, the presence of novel intrinsic absorbance and luminescence was observed in α_3C, α_3W, PRM, and HSA proteins. The effect of solvent parameters on ProCharTS and the possible influence of CT states on Trp fluorescence was investigated. The potential application of

ProCharTS to monitor protein unfolding and DNA-protein binding were also illustrated in the present thesis work.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

Page No

Chapter 2: Experimental techniques, Materials, and Methods

Chapter 3: Observing protein charge transfer spectra in proteins α_3C, α_3W, and PRM

CHAPTER 1

INTRODUCTION AND REVIEW OF LITERATURE

1.1 Intrinsic chromophores in proteins

Characterization of protein structure and function is fundamental to understanding key biological processes. Numerous techniques exist to study protein structure-function relationship, with some requiring the use of extrinsic chromophores and others employing sophisticated instruments/data analysis.[21,22] In this context, UV-Visible spectroscopy provides a convenient tool to characterize protein unfolding[6] and less frequently to monitor DNA-protein binding.[7] The intrinsic chromophore in proteins used for these studies are the aromatic amino acid residues. Proteins are expected to be optically silent in the parts of near UV-Visible electromagnetic spectrum where aromatic amino acid residues do not show absorbance and fluorescence.[23]

1.2 Conventional intrinsic chromophores in proteins

Intrinsic chromophores present in proteins are the peptide bonds, aromatic amino acid residues, associated prosthetic groups and coenzymes. The salient features of these chromophores are described below.

1.2.1 Peptide bonds

π-π* and n-π* electronic transitions happen in the peptide bond. Both transitions occur in the far UV region. The peptide bond contains π electrons as well as electrons in a non-bonding n-orbital. The π electrons are delocalized over the N, C, and O atoms whereas the non-bonded electrons are present at the O atom. Strong absorbance around 190 nm and another peak of lower intensity between 210—220 nm is observed from the peptide bond.[24] The intense peak has an ε_{max} of ~7000 $M^{-1}cm^{-1}$ and involves the π-π* transition whereas the less intense absorbance between 210—220 nm has an ε_{max} of ~100 $M^{-1}cm^{-1}$ involving the n-π* transition. The absorbance of peptide bond is dependent on conformation with the α-helical conformation of polyglutamic acid and polylysine showing less absorbance compared to β-sheet or random coil.[25] Peptide bond absorption at 220 nm has been used for protein estimation. The limitation with this approach is that most aromatic molecules and other amino acids absorb in this region too. Cys, Met, and His have significant absorbance in the far UV region.[25] His absorbs strongly in the 190—220 nm region but has negligible absorbance beyond 230 nm. Sulfur-containing cysteine, methionine, and cystine absorb in the far UV region with cystine absorbing even in the near UV region. These transitions happen without the involvement of delocalized π electrons. Cystine ε_{max} at 290 nm = 50 $M^{-1}cm^{-1}$ and contributes significantly to the absorbance of proteins rich in disulfide bonds.

1.2.2 Aromatic amino acids

Aromatic amino acids contribute to protein absorbance in the near UV region. π-π* transitions involving the π-electrons delocalized over the aromatic amino acid sidechain (Figure 1.1) gives rise to the observed absorbance.

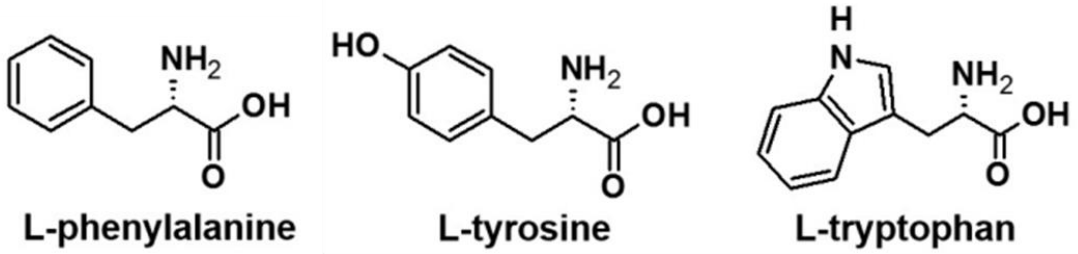


Figure 1.1: Structure of aromatic amino acids Trp, Tyr, and Phe (Adapted from Biophysical Chemistry, Part II, 1980).[8]

Trp contributes maximally to the near UV absorbance spectra of proteins because of high molar extinction coefficient. In the near UV region, Trp absorbance maximum is at 280 nm, Tyr has an absorbance maximum at 275 nm, and Phe at 257 nm (Figure 1.2). The absorbance of Trp extends till almost 320 nm and Tyr does not absorb beyond 295 nm. Trp ε_{max} = 5600 $M^{-1}cm^{-1}$ at 280 nm, Tyr ε_{max} = 1400 $M^{-1}cm^{-1}$ at 275 nm, and Phe ε_{max} = 200 $M^{-1}cm^{-1}$ at 257 nm.[8]

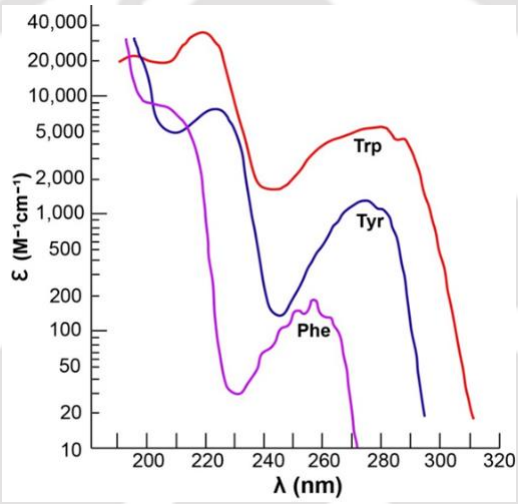


Figure 1.2: Absorption spectra of aromatic amino acids in the near and far UV region (Adapted from Biophysical Chemistry, Part II, 1980).[8]

The absorbance of aromatic amino acids is sensitive to its environment and is affected by solvent parameters and solvent exposure. Variation of pH affects the Trp and Phe absorbance moderately and affects the Tyr absorbance significantly. Alkaline pH causes a deprotonation of the –OH group in the Tyr side chain causing a red shift of the observed absorbance.[8] The difference in the absorbance spectra of native and unfolded protein is used to monitor protein unfolding.

The fluorescence spectra of Trp, Tyr, and Phe upon excitation at 266 nm are shown in Figure 1.3. Trp, Tyr, and Phe fluorescence are similar to the spectra of indole, phenol, and benzene solutions respectively. The fluorescence spectra of Phe is a mirror image of its absorption spectra, whereas the spectra of Tyr and Trp do not mirror their absorbance and appear broad and featureless. Phe, Tyr, and Trp fluorescence are significantly quenched at extreme pH.[26]

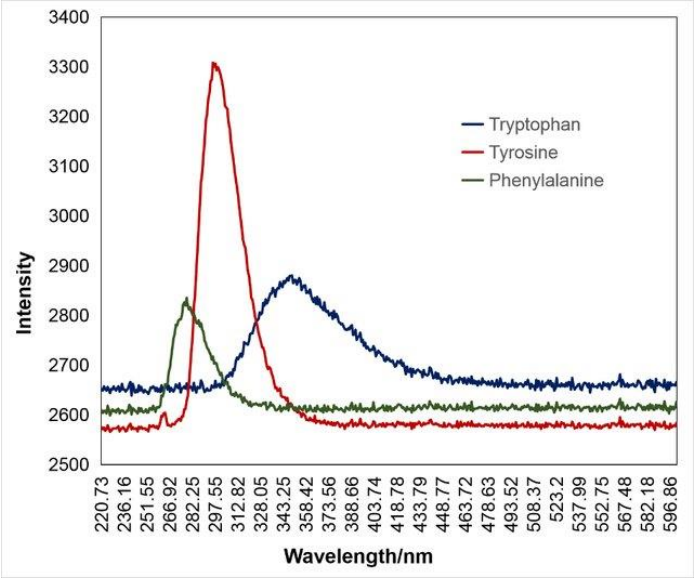


Figure 1.3: The fluorescence spectra of aromatic amino acids at neutral pH (deionized water) upon excitation at 266 nm (Adapted from Du *et al*, 2020).[27]

The parameters associated with aromatic amino acid fluorescence are reported in Figure 1.4. Quantum Yield (QY) of Phe is low but Tyr and Trp have moderate quantum yields. Tyr fluorescence is difficult to observe when Trp residues are present in the protein because of: (1) low Tyr extinction coefficient, (2) moderate quantum yield, and (3) energy transfer between Trp and Tyr complicating the observed emission.[26]

When all three aromatic amino acids are present, Trp is selectively excited using 295 nm wavelength. The sensitivity of Trp quantum yield and emission maximum to the environment has allowed the extensive use of Trp fluorescence in monitoring protein unfolding and protein-protein interactions.

Species	λ^{F}_{max} (nm)	λ^{ex} (nm)	Band width (nm)	Quantum yield	Lifetime (ns)
Phenylalanine	282	260	—	0.02	6.8
Tyrosine	304	275	34	0.14	3.6
Tryptophan	353	295	60	0.13	3.1 (mean)

Figure 1.4: The emission max λ^{F}_{max}, quantum yield, and fluorescence lifetime of aromatic amino acids at neutral pH (deionized water) (Adapted from ultraviolet spectroscopy of proteins, 1987).[26]

1.2.3 Prosthetic groups and coenzymes

The prosthetic groups and coenzymes associated with proteins contribute to absorbance in the near UV-Visible region. Prosthetic groups absorb light in the visible range and are associated with important photosensory biological functions. Absorbance originating from tightly (prosthetic groups) or loosely bound (coenzymes) non-protein components can involve charge transfer (CT), d-d transitions, pi, or non-bonded electrons.

FAD, NADH, and NAD^+ are important coenzymes which show spectra in the UV-Vis region. These coenzymes absorb due to the presence of aromatic groups in the structure. FAD has maximum absorbance at 450 nm (ε_{max} = 11,300 $M^{-1}cm^{-1}$), NADH at 340 nm (ε_{max} = 6,220 $M^{-1}cm^{-1}$), and NAD^+ at 259 nm (ε_{max} = 16,900 $M^{-1}cm^{-1}$).[28] The absorbance maximum of porphyrin ring in heme is at 404 nm (ε_{max} = 170,000 $M^{-1}cm^{-1}$).[29] These compounds absorb strongly in the visible region which is distinct from peptide bond and aromatic amino acid absorbance.

1.3 Novel intrinsic chromophores

Proteins devoid of aromatic amino acid residues and lacking prosthetic groups are expected to be optically silent in the near UV-Visible region as protein absorbance and luminescence is conventionally expected to originate from aromatic amino acid residues Phe, Tyr, and Trp.[23–25] Even in proteins containing aromatic amino acid residues, absorbance beyond 325 nm is not expected.[25] However, multiple recent reports have highlighted protein absorbance and luminescence in regions of the electromagnetic spectrum where proteins are not expected to show any optical activity.[9,10] These reports date back to 2001[12] but computational insight and potential use as a label-free detection technique has renewed interest in understanding the origin and potential applications of the novel spectra. The origin of the spectra has been the discussion of many reports and various phenomena of charge transfer and recombination,[9,15] carbonyl group based autofluorescence,[11] and proton transfer[10] have been suggested to account for this novel absorbance and luminescence, and will also be the subject of the present discussion.

1.3.1 Protein Charge Transfer Spectra

The involvement of electron transfer (ET) in key biological processes such as respiration and photosynthesis is a known phenomenon.[30] The presence of long-range ET has been shown in multiple proteins such as cytochrome c, haemoglobin, myoglobin, and azurin.[31] Cytochrome c heme group is involved in accepting electrons from cytochrome bc1 complex and subsequent transfer to complex IV. Haemoglobin and myoglobin contain the O_2-binding and iron-containing heme prosthetic group. In myoglobin, charge transfer involves the O_2 and the Fe at the metal centre.

Azurin is a bacterial copper-containing protein on which extensive electron transfer studies have been performed.[32] Charge transfer from the Cys thiol ligand to Cu^{2+} gives azurin an intense blue colour. Intramolecular electron transfer in azurin to intrinsic and extrinsic (covalently attached) electron transfer partners has given significant insight into photoinduced electron transfer in proteins. The influence of ET distance, secondary structure of the bridging peptide residues, and donor-acceptor structure on electron transfer have been characterized with azurin as a model system.

Two models have been proposed to explain photoinduced electron transfer in proteins: (1) superexchange model and (2) hopping model.[33] The electron transfer from donor to acceptor occurs in a single step according to superexchange model. According to the hopping model, ET

involves multiple steps which incorporate electron hopping via aromatic amino acid sidechains. Thus, rather than ET occurring in a single slow step, ET involves multiple short but faster steps. The enzyme ribonucleotide reductase involves the long distance electron transfer from Tyr-122 to Cys-439 residue.[33] The distance between these two residues is 35 Å and the role of Tyr-356, Tyr-730, and Tyr-731 as relay amino acid residues has been ascertained. Tyr sidechain facilitates electron transfer by formation of the Tyrosyl radical and electron hopping is thought to occur only by the aromatic amino acid sidechains. Both the superexchange and hopping models are used to explain the mechanism of long-distance photoinduced electron transfer in proteins.

Both types of charge transfer (CT) transitions (electron and proton transfer) have been shown in proteins. The most prominent example of excited-state proton transfer (ESPT) with Glu-222 as the proton acceptor was elucidated in the Green Fluorescence Protein.[34] The photoinduced conversion of GFP from neutral to anionic form involves ESPT to E222 with both the forms having distinct photophysical properties.

Protein Charge Transfer has been hypothesized as the possible reason for protein α_3C strongly absorbing light in the near UV-Visible region with a long tail extending till 800 nm.[9] The background work for this study was a report of novel absorbance and luminescence from concentrated L-Lysine solutions made in 2001 by Swaminathan *et al*.[12] Significant absorbance was observed at 270 nm (ε = 0.34 $M^{-1}cm^{-1}$) and novel luminescence from Lys was also observed upon excitation at 355 nm giving a peak at ~435 nm. In a follow up study reported in 2004, absorbance of Human Serum Albumin (HSA) beyond 325 nm was attributed to abundance of Lys residues in close proximity.[13] Significant absorbance by HSA beyond 325 nm (in the region where aromatic amino acids are not expected to absorb) was observed but not from control proteins Subtilisin Carlsberg, Lysozyme, and Barstar. The importance of Lys-Lys contacts in the generation of novel absorbance was ascertained as the number of these contacts were much higher in the PDB structure of HSA compared to Lys contacts in the control proteins. The possible role of protein fold in the HSA absorbance (ε_{325} = 1546 $M^{-1}cm^{-1}$) being much higher than Lys was ascertained in a 2017 study by Prasad *et al*.[9]

Charged amino acid solutions absorbed strongly in the 250—400 nm region compared to negligible absorbance shown by uncharged amino acids (Figure 1.5).[9] Absorbance by concentrated solutions of charged amino acids (1 M) hinted towards the possibility that not only Lys-Lys contacts, but other charged amino acid contacts may lead to strong unexpected absorbance in proteins.

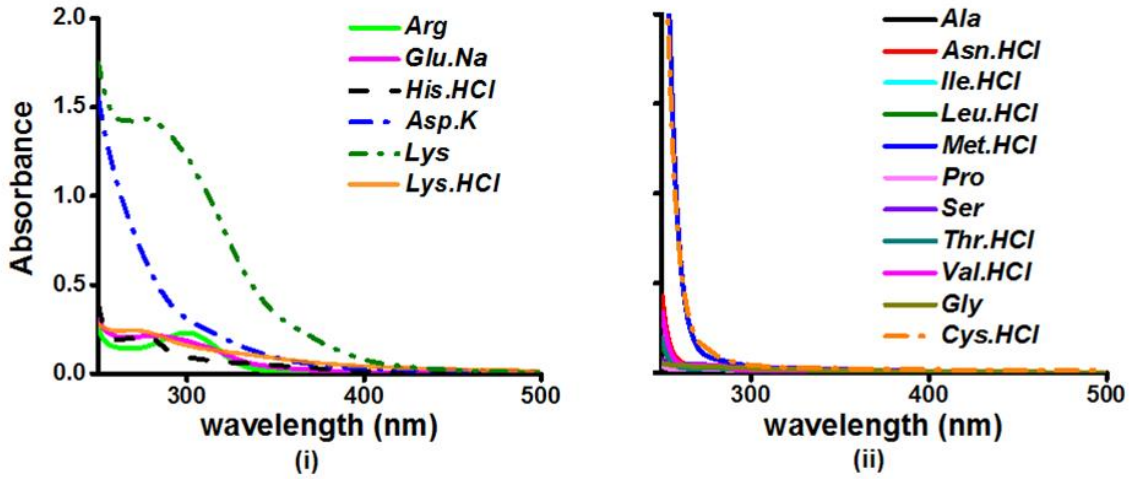


Figure 1.5: Absorbance spectra of charged amino acids (i) compared with uncharged amino acids (ii). The concentration of the amino acids used was 1 M and solutions were made in deionized water (Adapted from Prasad *et al*, 2017).[9]

α_3C is a synthetic protein devoid of aromatic amino acid residues and containing an abundance of charged amino acid residues. Of the 67 constituent amino acid residues there are 17 Lys, 17 Glu, and 2 Arg. Figure 1.6 shows the reported novel α_3C absorbance in the near UV-Visible region. The observed absorbance at three representative wavelengths was shown to increase linearly with concentration in accordance with the Beer-Lambert's law. The linear increase in α_3C absorbance with increasing concentration (inset of Figure 1.6) ruled out the possibility of protein aggregation at higher concentration.

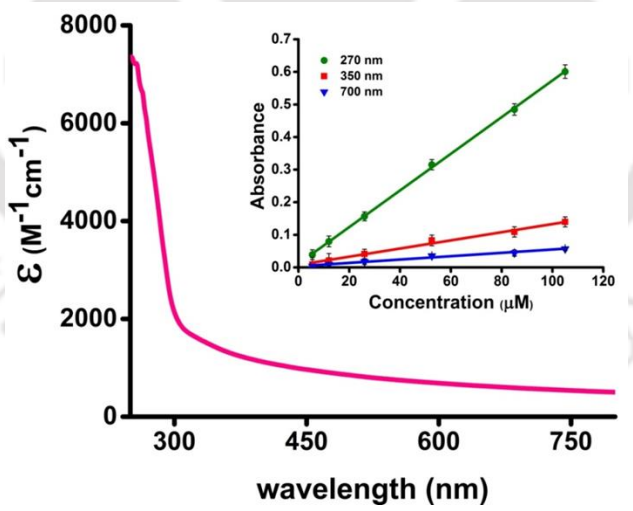


Figure 1.6: Protein Charge Transfer Spectra in α_3C and inset showing linear increase of absorbance with concentration (Adapted from Prasad *et al*, 2017).[9]

1.3.2 Computational studies on charged amino acid sidechain contacts in proteins

A combination of Molecular Dynamics (MD) simulations and Time-Dependent Density-Functional Theory (TDDFT) calculations were done on α_3C.[9]

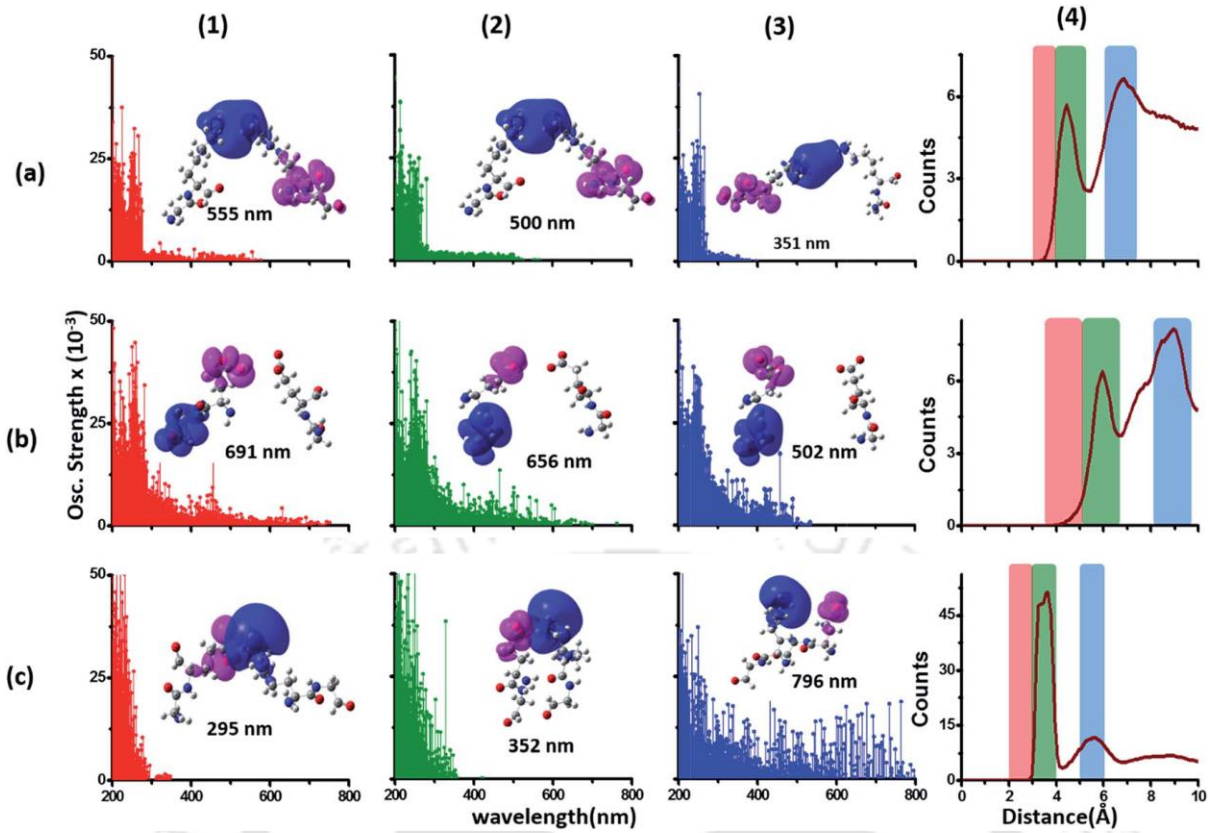


Figure 1.7: Computed absorption spectra for Lys-Lys (a1-3), Glu-Glu (b1-3), Lys-Glu dimers (c1-3), and radial distribution function (RDF) plots of Glu C_C and Lys N_A atom pairs (4a-c). Low energy CT transitions well into the visible region were observed for Lys-Lys contacts at close distance (a1), Glu-Glu contacts at close distance (b1), and Lys-Glu contacts separated from each other (c3). The RDF of Lys-Lys (4a), Glu-Glu (4b), and Lys-Glu pairwise distance obtained from the MD simulation run of α_3C is also shown (Adapted from Prasad *et al*, 2017).[9]

Using TDDFT, the theoretical absorption spectra of Lys-Lys, Glu-Glu, and Lys-Glu dimers extracted from snapshots of α_3C MD simulation run were calculated (Figure 1.7). Lys-Lys and Glu-Glu close contacts (Lys-Lys pairs 3—4 Å apart and Glu-Glu pairs 3.5—5 Å apart), and Lys-Glu weak contacts (5—6 Å apart) were found to be involved in the generation of the observed α_3C absorbance. Specifically, it was reported that charge transfer from peptide backbone to Lys sidechain in Lys-Lys close contacts (Figure 1.7 a1), from Glu sidechain to peptide backbone in Glu-Glu close contacts (Figure 1.7 b1), and from Glu to Lys sidechain in weak Lys-Glu contacts (Figure 1.7 c3) was occurring.[9]

Essentially a combination of these Peptide backbone to Sidechain (PBS) and Sidechain to Sidechain (SS) Charge Transfer (CT) transitions generated the observed α_3C absorbance. PBS-CT, SS-CT, and the donor-acceptor groups involved in charge transfer are illustrated in Figure 1.8 (b) and (c). The observed spectra originated due to charge transfer in proteins and was thus named Protein Charge Transfer Spectra (ProCharTS).

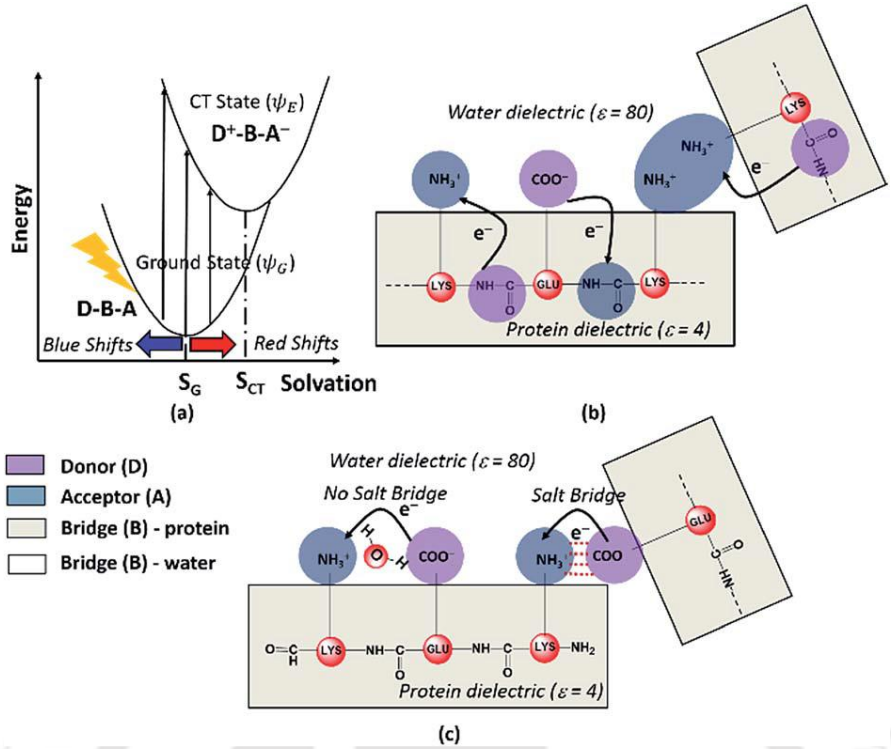


Figure 1.8: Model depicting photoinduced charge transfer in proteins with a Donor (D), Bridge (B), and acceptor (A) system. (a) S_G and S_{CT} correspond to the minimum energy configurations of the ground and excited states, showing that upon photoexcitation charge transfer will happen from donor to the acceptor. (b) PBS-CT from peptide backbone to Lys sidechain, Glu sidechain to peptide backbone, and peptide backbone to Lys-Lys pair in close proximity. (c) SS-CT from Glu to Lys in Lys-Glu weak contact (Adapted from Prasad *et al*, 2017).[9]

The first report of CT transitions in a protein devoid of metal centres or aromatic groups was thus made in α_3C which contains an abundance of Lys and Glu. In a subsequent computational study, CT transitions upon excitation with wavelengths longer than 250 nm were reported for all the charged amino acid residues and naturally phosphorylated amino acid residues.[14] PBS-CT was thus additionally observed in Arg, Asp, His, and the phosphorylated form of Ser, Thr, and Tyr. The profile and intensity of the computed spectra depended on the nature of the donor, acceptor, and bridge group (donor, bridge, and acceptor groups are illustrated in Figure 1.9) and was also sensitive to the type of secondary structure from which the monomer used for calculation was extracted. This computational study thus predicted the possibility of charge transfer in proteins abundant in any of the charged amino acid residues.

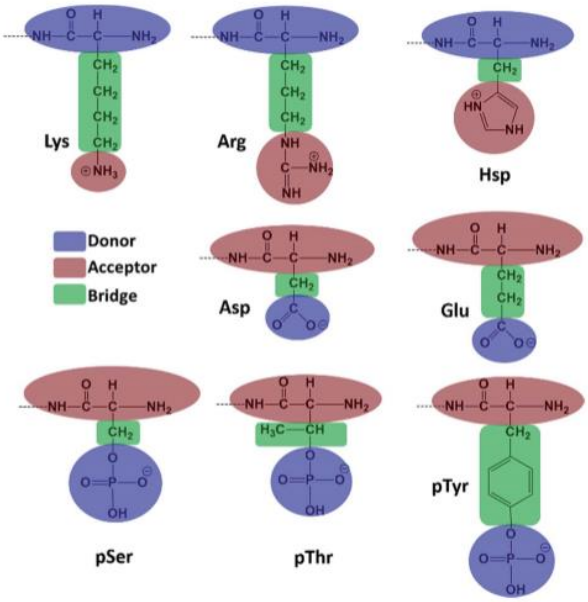


Figure 1.9: PBS-CT expected in charged amino acid residues and phosphorylated amino acid residues with their donor, acceptor, and bridge groups highlighted (Adapted from Mandal *et al*, 2018).[14]

Human c-Myc PEST Wt, Trp mutant PEST M1, and Dehydrin 1 (DHN1) are intrinsically disordered proteins with abundant and diverse charge content. The presence of ProCharTS in these proteins was reported (Figure 1.10). Novel absorbance observed from three proteins with distinct charge compositions (Glu, Asp, and His rich PEST proteins, only 1 Lys and 0 Arg in PEST proteins; presence of all 5 charged amino acid residues in DHN1) validated that charged amino acid residues of all types could generate ProCharTS albeit with different intensities.[35]

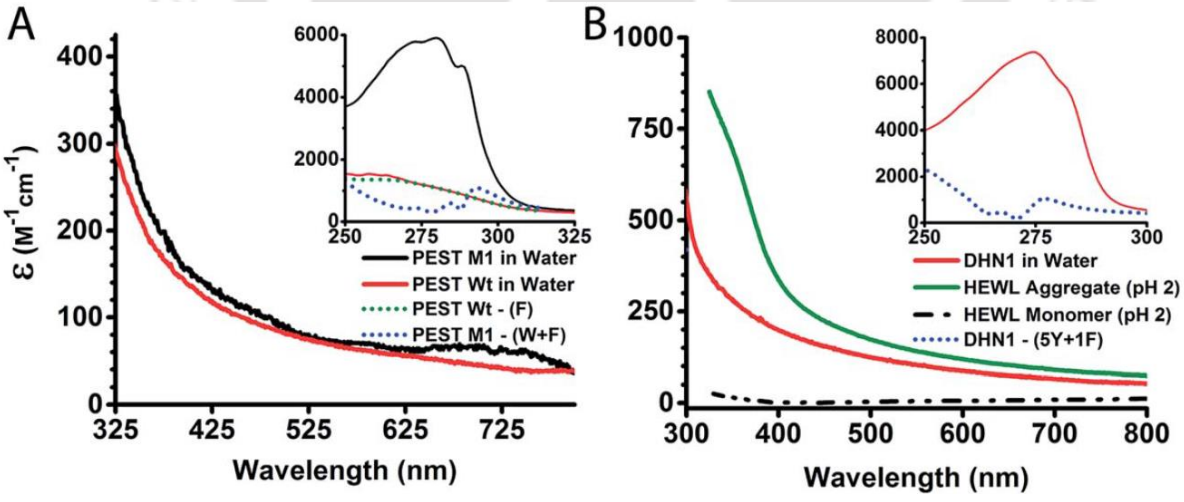


Figure 1.10: ProCharTS in PEST, DHN1 monomeric proteins, and HEWL aggregates (Adapted from Ansari *et al*, 2018).[35]

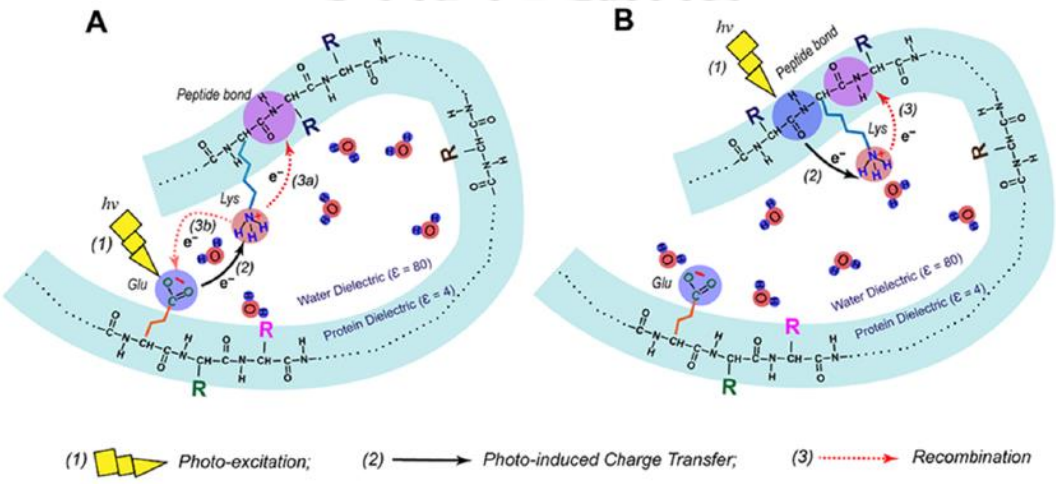
ProCharTS depends on electrostatic interactions between charged amino acid sidechains which depend on protein fold and solvent parameters such as pH, temperature, and the presence of salt. Ansari *et al* characterized how solvent parameters affected ProCharTS intensity. PEST Wt and M1

are intrinsically disordered but gain structure at pH 3 as evidenced by CD spectra. The structure gain led to a substantial increase in PEST ProCharTS at this pH, possibly due to increased charged amino acid sidechain contact formation upon protein folding. It is expected that salt counterions would screen protein sidechain charges and disrupt electrostatic contacts. α_3C ProCharTS absorbance intensity reduced in the presence of 250 mM NaCl correlating well with an expected decrease of electrostatic contacts in the presence of salt. Hen egg white Lysozyme (HEWL) aggregates showed a several-fold increase in ProCharTS compared to monomeric HEWL due to the process of HEWL aggregation resulting in the substantial increase in intermolecular charged amino acid contact formation (Figure 1.10B).

It has thus been shown that several structured and unstructured proteins having high charge content and no/little aromatic amino acid residues show strong absorbance in the near UV-Visible region not associated with any known chromophores in proteins. The intensity of the novel absorbance observed depends on the nature and number of charged amino acid sidechain contacts in the protein. The involvement of photoinduced PBS-CT and SS-CT in the origin of ProCharTS was uncovered. The effect of solvent parameters on ProCharTS was characterized in multiple proteins and a correlation between protein fold and ProCharTS intensity also illustrated in these cases.

1.3.3 Charge recombination luminescence in proteins

Numerous monomeric proteins abundant in charged residues were reported to show luminescence upon excitation with near UV wavelengths in the 325—400 nm region.[15] The proteins studied included natively folded proteins such as HSA and α_3C; and unstructured proteins such as human c-Myc PEST fragments and DHN1. The features of the observed luminescence included a similarity in steady-state emission and excitation profile, moderate to low quantum yields, and high Stokes shifts. The slope of integrated luminescence intensity increase with concentration correlated with the abundance of charged amino acid residue contacts in the 3-dimensional protein fold. Similarity in the mean luminescence lifetime of proteins (0.4—2.9 ns) and multi-exponential decay kinetics was also observed. The observed similarity in photophysical features of luminescence from multiple proteins hinted towards a common origin. The mechanism of charge recombination luminescence was hypothesized to be involved in the novel luminescence.



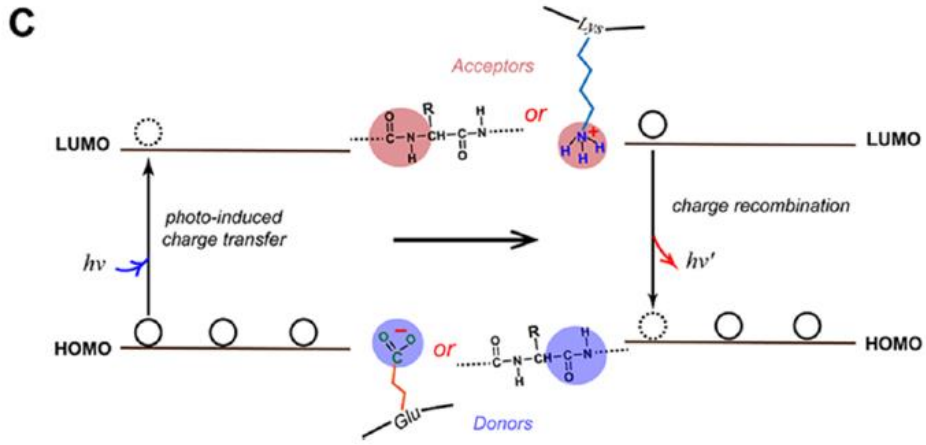


Figure 1.11: Schematic illustrating charge recombination luminescence (Adapted from Kumar *et al*, 2020).[15]

(1) Photo-excitation, (2) Photo-induced Charge Transfer, and (3) recombination could lead to the generation of charge recombination luminescence (Figure 1.11). Upon photo-excitation, charge transfer can happen either from the donor Glu sidechain/peptide bond to the acceptor peptide bond/Lys sidechain. Subsequently, charge recombination within the same HOMO-LUMO pair or a proximal HOMO-LUMO can occur, with the efficiency of this recombination determining the quantum yield. It was hypothesized that the presence of multiple proximal charged amino acid sidechains in HSA facilitated greater charge recombination resulting in higher quantum yield.

1.3.4 The effect of ProCharTS on Trp fluorescence

Trp fluorescence decay fits to a sum of two exponentials (short component of ~0.5 ns and long component of ~3.1 ns) in deionized water, whereas the fluorescence decay of N-Acetyl L-Tryptophanamide (NATA) is single exponential. Additionally, single Trp containing proteins have multi-exponential fluorescence decays which has been attributed to Trp rotamers.[36] However, the exact reason for the multi-exponential decay of Trp fluorescence is not yet ascertained.

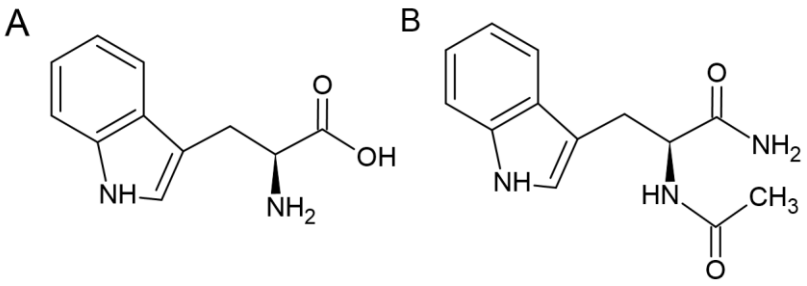


Figure 1.12: Structure of (A) Tryptophan and (B) Trp analogue NATA.

Due to significant overlap between Trp absorbance and fluorescence with ProCharTS absorbance and luminescence, the additive contribution of ProCharTS luminescence to Trp fluorescence was ascertained by our group recently.[16] Energy transfer between CT states and Trp was not observed. The deviation of NATA fluorescence decay from single exponential to sum of two exponentials

in the presence of charged amino acid residues and proteins rich in charged amino acid residues was shown (Figure 1.13).

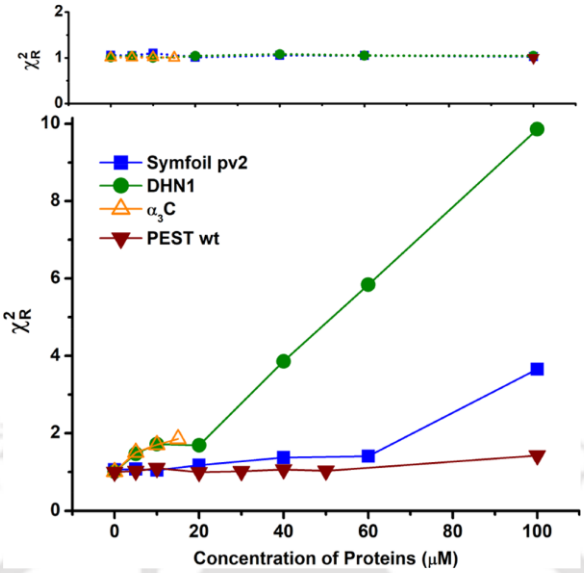


Figure 1.13: Deviation of NATA fluorescence decay from single (bottom) to sum of two exponential fits (top) in the presence of proteins rich in charged amino acids (Symfoil pv2, DHN1, α3C, and PEST wt) (Adapted from Kumar *et al*, 2022).[16]

1.3.5 Aggregation induced emission

Crystals of proteins with varying amino acid compositions (γ-B crystallin, HEWL, and 6 other proteins) were shown to emit blue radiation upon excitation with near UV wavelengths.[37] The observed emission was found to be independent of the presence of aromatic amino acid residues as a protein crystal derived from a γ-crystallin 14-mer (containing no aromatic amino acid residues) also showed similar the weak novel emission. The emission observed was photo-bleachable and quenchable by potassium iodide (KI) confirming that the emission obtained was not due to Raman scattering. Based on these observations, the authors proposed the role of peptide bond electrons, delocalized by the formation of intramolecular and intermolecular H-bonds in generation of the observed Aggregation induced emission (AIE).

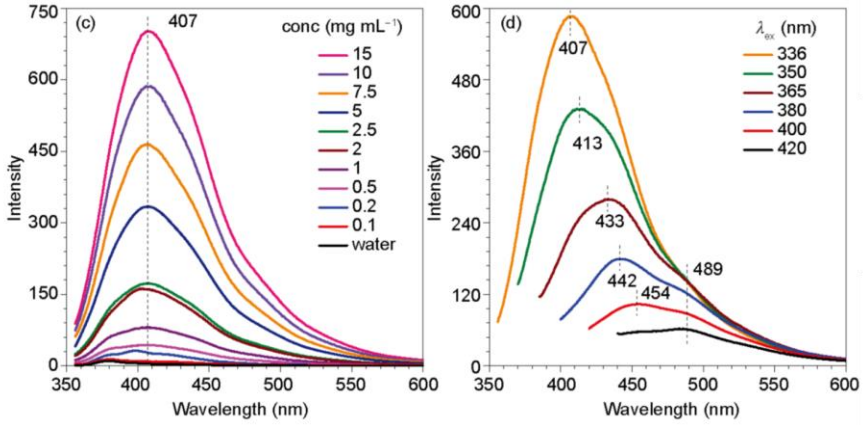


Figure 1.14: Emission spectrum of Poly-L-Lysine at λ_{ex} = 336 nm with increasing concentration (left) and 15 mg/mL Poly-L-Lysine with multiple λ_{ex} (right) (Adapted from

Chen *et al*, 2018).[38]

In a study by Chen *et al*, a poly-amino acid crystal (Figure 1.14) and eight non-aromatic amino acid crystals showed novel intrinsic emission.[38] The excitation and emission profiles matched with similar observations made in previous reports, however the mechanism of Clustering Triggered Emission (CTE) was proposed for the origin of the luminescence. Based on L-Ser crystal structure, it was proposed that in amino acid crystals amino, hydroxyl, and carbonyl groups can come close together, allowing their electronic clouds to overlap and delocalize. These delocalized electron clouds would correspond to low energy electronic transitions, giving rise to emission in the visible region upon excitation with near UV wavelengths leading to the observed phenomenon of aggregation induced emission.

The prevalent intrinsic blue emission upon near UV excitation from fibrils of 5 amyloidogenic proteins ($A\beta_{42}$, $A\beta_{40}$, $A\beta_{33\text{-}42}$, K18 Tau, and I59T lysozyme) has also been reported.[39] Similarity in the excitation-emission profile and fluorescence lifetimes of the fibril luminescence was also observed by Chan *et al*.[39] An increase in the intrinsic luminescence with protein aggregation was observed. The potential application of the novel luminescence for LABEL-FREE detection of protein aggregation was proposed (Figure 1.16). Proton transfer in the amyloid fibril H-bond networks was cited as a possible explanation for the observed luminescence although computational studies into electronic structures involved in the luminescence was not performed.

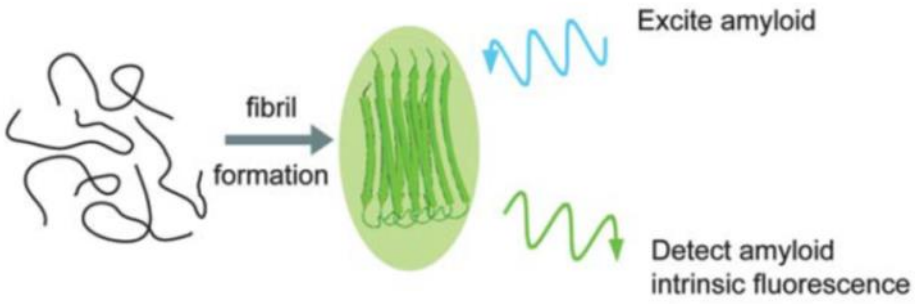


Figure 1.16: Illustration of intrinsic fluorescence from amyloid fibrils (Adapted from Chan *et al*, 2013).[39]

Computational studies on fibrils of amyloid-beta ($A\beta_{1\text{-}42}$) derived peptides revealed the presence of proton transfer across the N and C terminus of amyloid fibrils.[10] Adjacent protofilaments in amyloid fibrils are bonded to each other by a dense network of H-bonds. This network was shown computationally to be a region of delocalized protons which can facilitate low energy optical excitations explaining the observation of the blue luminescence from amyloid fibrils. Figure 1.15 shows the excitation and emission spectra of $A\beta_{1\text{-}42}$ amyloid fibrils and the effect of extreme pH on measured novel luminescence. The decrease in luminescence at extreme pH was explained by reduced proton transfer at pH 0 and pH 14. At pH 0 all ionizable groups at the termini should be protonated, whereas at pH 14, the ionizable groups should be fully deprotonated, in both conditions greatly reducing the proton transfer across the N- and C- termini. Hence with a combination of

computational and spectroscopic studies, the contribution of proton transfer in dense H-bonded networks to the photophysical properties of amyloid fibrils was discovered.

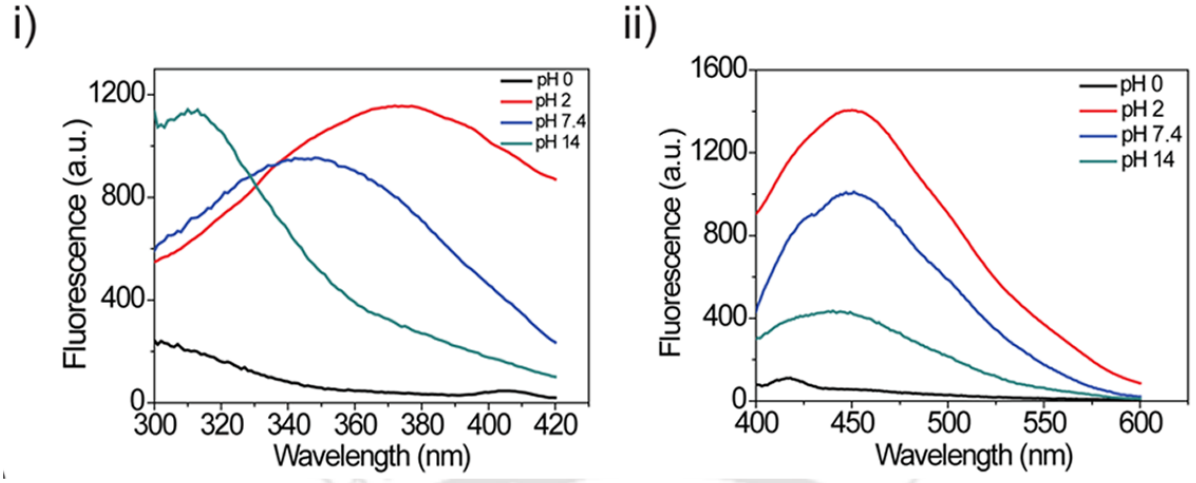


Figure 1.15: The effect of pH on Aβ$_{1-42}$ amyloid fibril luminescence (i) excitation and (ii) emission spectra (Adapted from Pinotsi *et al*, 2016).[10]

The intrinsic luminescence from concentrated HCl and NaOH solutions was reported recently.[40] HCl and NaOH luminescence appear as broad peaks with emission maximum at ~300 nm and ~400 nm respectively. NaOH luminescence can overlap with the reported blue luminescence from proteins as HCl and NaOH are occasionally used in buffer preparation. Thus the contribution from HCl and NaOH needs to be duly subtracted from protein luminescence.

The deep blue luminescence in amyloid fibrils derived from 13 amyloidogenic proteins was characterized.[41] The first report of near-infrared luminescence from amyloid fibrils (Het-s prion domain emission shown in Figure 1.17), along with deep blue luminescence, was made by Pansieri *et al*. The discovery of near-infrared luminescence is quite remarkable and mechanistic insight into this aspect of amyloid fibril photochemistry is eagerly awaited as it provides an additional spectroscopic window for the label-free detection of pathogenic amyloid fibrils.

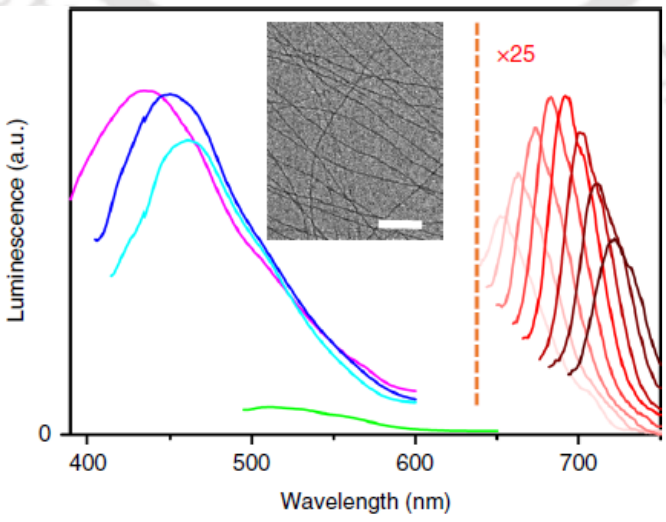


Figure 1.17: Het-s prion domain amyloid fibrils show novel deep blue luminescence and near-infrared emission (Adapted from Pansieri *et al*, 2019).[41]

Circular dichroism of concentrated γ-B crystallin solutions (100 mg/mL) in the 350—550 nm region was reported.[42] Circular dichroism in this region is not expected from the peptide bond and was attributed to the novel chromophore previously reported in γ-B crystallin.[37]

The involvement of CT in absorbance and luminescence observed from amyloid fibrils was shown by Hassanali *et al*.[43] Spectroscopic measurements and joint computational studies were performed on fibrils prepared from Aβ derived oligopeptides. TDDFT calculations were done on the three possible dimer configurations of a protofilament shown in Figure 1.18 (left): linear parallel (a), anti-parallel (b), and parallel (c). Charge transfer and the sensitivity of CT transitions to conformation was observed at the dimer termini. The direction, oscillator strength, and location of CT corresponding to the lowest energy excitation is shown in Figure 1.18 (right) for each dimer configuration. The electron and hole situated on the N- and C-terminus are shown in red and blue respectively, along with the direction of CT. Thus CT transitions have also been reported at the termini of amyloid fibrils.

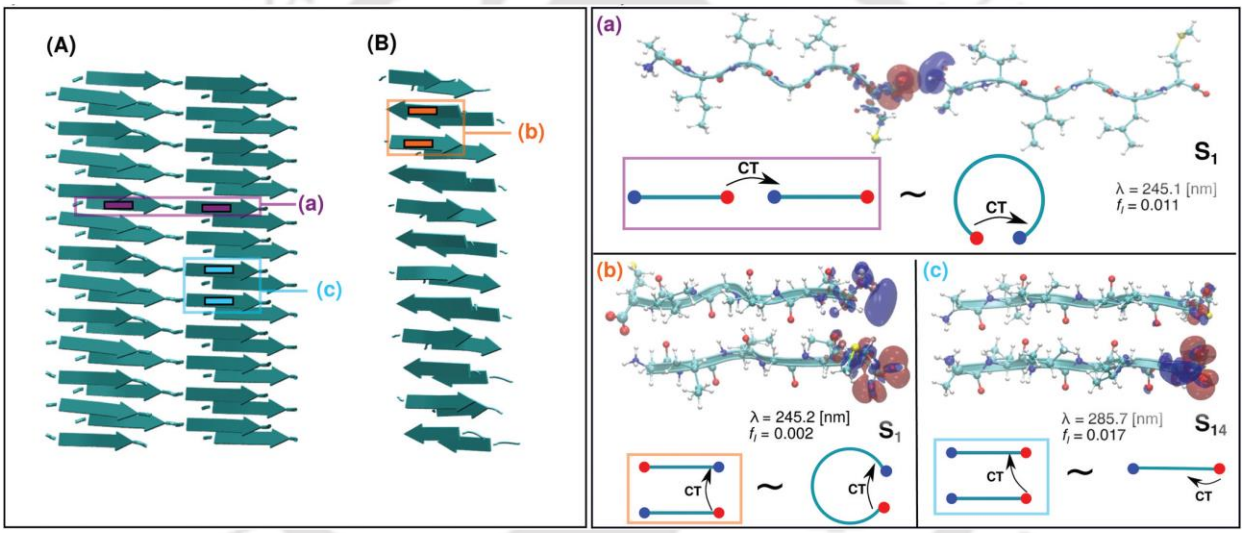


Figure 1.18: Charge transfer at the termini of amyloid fibrils formed from Aβ derived oligopeptides (Adapted from Hassanali *et al*, 2019).[43]

Blue luminescence and near-infrared luminescence from amyloid fibrils of multiple proteins have thus been reported. Similarity in the excitation and emission spectra of many fibrils has been elucidated. Circular dichroism of the novel chromophore in a protein was also shown. Mechanisms as diverse as Clustering Triggered emission, Peptide bond electron delocalization and Proton-transfer have been proposed to explain the observed aggregation induced emission. Hence these reports collectively shed light on the novel photophysical properties of amyloid fibrils.

1.3.6 Carbonyl group based blue autofluorescence

Novel absorbance and fluorescence in concentrated solutions of monomeric HEWL, Bovine Serum Albumin (BSA), and β-Lactogluobulin (BLG) were reported in a 2017 study.[11] The proteins were shown to absorb wavelengths in the near UV region and emit in the deep blue region. The observed phenomenon was termed deep-blue Autofluorescence (dbAF) and the authors found the

presence of dbAF in multiple monomeric proteins. The authors systematically ruled out the presence of oligomeric forms and protein aggregates in the samples on which fluorescence measurements were made. Dynamic Light Scattering (DLS) experiments ruled out the presence of large protein aggregates and a linear increase in HEWL and BLG dbAF with concentration pointed to the absence of oligomeric species.

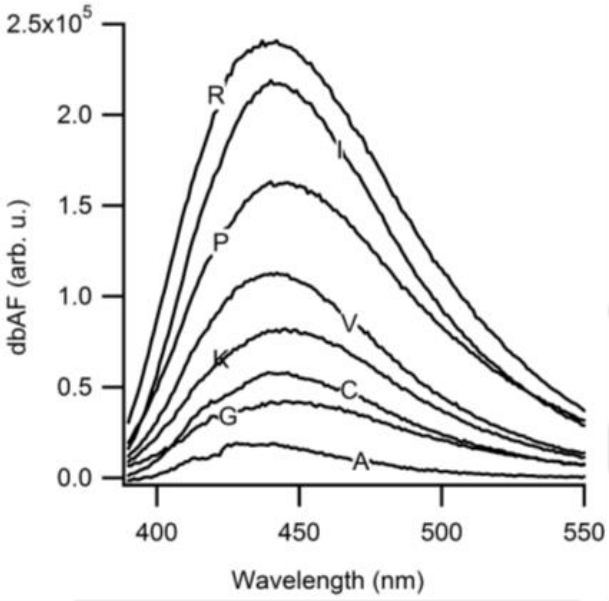


Figure 1.19: dbAF of multiple concentrated amino acid solutions. 20 mg/mL amino acid samples were made in 25 mM phosphate buffer at pH 2. The amino acids used included charged, hydrophobic, and uncharged amino acids (Adapted from Niyangoda *et al*, 2017).[11]

Eight amino acids of different types showed dbAF (Figure 1.19) with the similarity in emission profile (shape of emission curve and λ_{max} in the narrow range of 430—450 nm), suggesting a common origin. Based on the presence of dbAF in non-aromatic amino acid solutions, and the presence of intense fluorescence with similarity to dbAF from organic solvents containing carbonyl groups (acetone and formaldehyde), the authors concluded that the dbAF originated in the carbonyl group. The novel emission was absent from solvents without carbonyl groups (methanol and isopropanol).

In a 2018 study, based on a substantial increase in amino acid dbAF upon heating and the observation of oxidized amino acid products in the heated samples (with much greater autofluorescence) by ultrahigh-resolution mass spectrometry, a possible role of oxidization of amino acids in protein dbAF was also uncovered.[44]

Insulin dbAF was used to follow aggregation kinetics (Figure 1.20 A) to find out parameters such as fibrillization midpoint and fibril growth rate, which were in good agreement with conventional Thioflavin T (ThT) fluorescence assays and light scattering.[45] The dissociation of preformed insulin fibrils by denaturant guanidinium isothiocyanate (GdSCN) was followed by dbAF (Figure 1.20 B) to identify the denaturation midpoint. At high GdSCN concentration, a complete absence of insulin fibrils was expected and verified by ThT fluorescence. Residual fluorescence after

dissociation of fibrils at high denaturant conditions can only be concluded to come from monomeric insulin which was further proof that the novel deep-blue autofluorescence does not originate wholly due to protein aggregation.

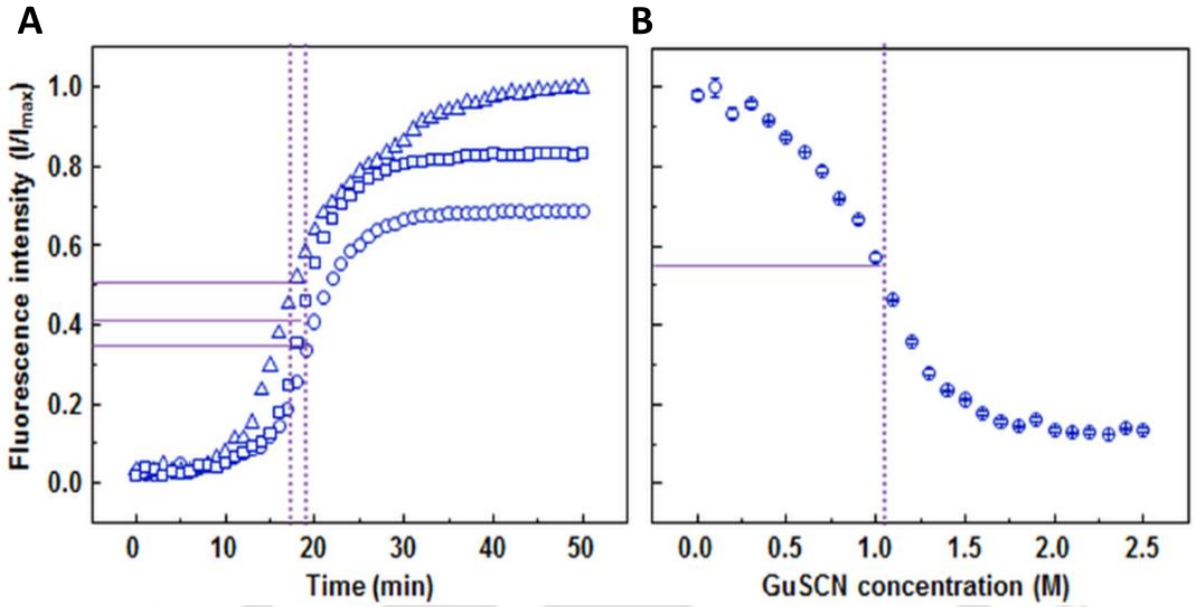


Figure 1.20: Insulin aggregation and dissociation studied by dbAF (Adapted from Ziaunys *et al*, 2019).[45] The increase in dbAF (experiment performed for triplicates) with protein aggregation was used to monitor seeded Insulin aggregation (Figure 1.20A). The dissociation of insulin fibrils by 6 M guanidine thiocyanate was monitored by dbAF (Figure 1.20B).

Systematic computational studies on the electronic structures involved in the origin of dbAF have not been done. However, it is clear that protein aggregation is not required for the observation of luminescence. The observation of absorbance and luminescence in monomeric proteins can be explained by protein charge transfer, based on the TDDFT studies done to predict the presence of charge transfer in proteins and corresponding spectroscopic measurements done in model protein α_3C.

1.4 Protein folding

1.4.1 Protein folding process

The complex process by which a protein synthesized as a random coil folds and attains its native functional state is an important research problem in protein biophysics. Protein structure is classified hierarchically into the primary, secondary, tertiary, and quaternary structures. Proteins are synthesized using their mRNA template by the ribosome as a linear chain of amino acid residues. The amino acids and more importantly the sequence in which they are present (primary structure) determine which amino acid residues will come close together in the final three-dimensional structure. In the next level of organization, amino acid residues form intramolecular H-bonds between amide hydrogens and carbonyl oxygen of the peptide backbone to form secondary structure elements. The formation of secondary structures like alpha-helices and beta-sheets is one of the earliest steps in protein folding. Subsequently, these secondary structures come

together in a way that hydrophobic portions are buried in the core and hydrophilic portions remain exposed to the solvent, thus attaining the tertiary structure of the protein. The tertiary structure of a protein is comprised of one polypeptide chain, but multiple such chains can come together to form the quaternary structure.[46]

An energy landscape of protein folding represents all the possible pathways a nascent polypeptide can take to attain the native state.[47] Protein folding happens *in vitro* in the ms timescale but folding can take up to seconds and can also happen within a few μs. At the beginning of the folding process, there exists an ensemble of denatured conformations which search through the conformational space to attain the native state. The first step in protein folding is the emergence of contacts between key local residues in domains. Protein folding happens in a modular process and the above local interactions lead to secondary and tertiary structure formation. The formation of the transition state in which intra- and inter- domain native-like interactions exist (shown in yellow in Figure 1.21) is a crucial step. It is a state which all unfolded proteins must attain before they become fully folded. The molecular evolution of proteins has happened in a way that rather than exploring the entire conformational space accessible to a folding protein, the folding is funneled to proceed via intermediates in which native-like interactions dominate.

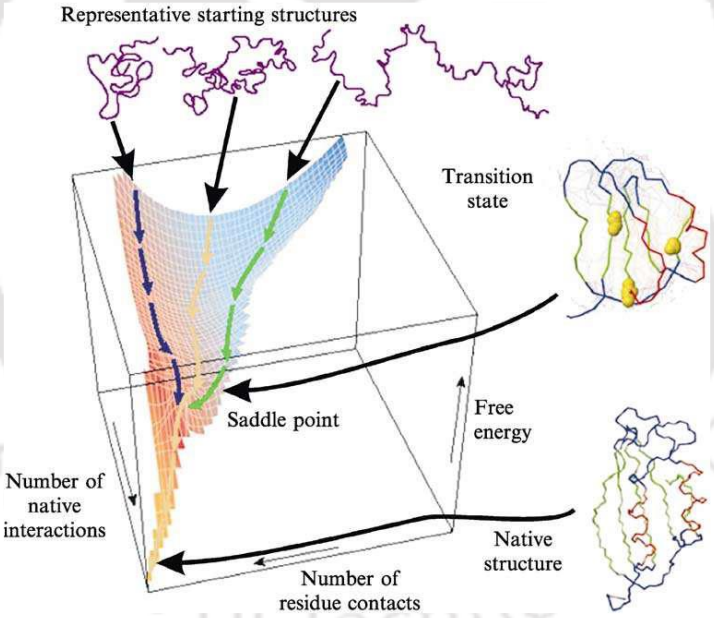


Figure 1.21: Illustration of protein folding using energy landscape diagram. The energy landscape depicts the crucial points in the attainment of native structure including the transition state and increase in number of native interactions with the protein folding process (Adapted from Dyson, 2003).[47]

The characterization of protein folding deals with studying the structures present in the native state of the protein as well as the kinetics of their formation. A combination of equilibrium and non-equilibrium studies using ensemble, single-molecule, and computational methods has led to significant advances in understanding the protein folding problem.

Early protein folding experiments dealt with characterization of the dominant forces involved in the process. In a seminal review in 1990, Ken Dill summarized the types of interactions and their relative contribution to the native protein structure. Among electrostatic interactions, H-bonding, van der Waal, and hydrophobic interactions, maximum contribution to protein stability was derived from the weak but numerous hydrophobic interactions.[48] The presence of hydrophobic amino acid residues at the core of solved crystal structures was observed. Burial of hydrophobic amino acid residues in protein core maximizes the hydrophobic interactions and solvent entropy, thus driving a random coil protein to attain its native three-dimensional structure. Computational studies additionally showed that incorrect folding of proteins was correlated with non-polar amino acid sidechains getting exposed to the solvent and that hydrophobic interactions were the principal driving force behind protein folding.

Protein stability depends on solvent conditions such as pH, salt concentration, temperature, and the presence of denaturants.[49] At pH significantly different from the pI, proteins unfold because of an increase in the net charge at the protein surface causing electrostatic repulsion. Similarly, protein denaturation at high temperatures is caused by a gain in conformational entropy. Interestingly proteins denature at cold temperatures as well, driven by the increase in the hydration of hydrophobic amino acid residues in the protein core with reducing temperature. The presence of denaturants such as urea and GdnHCl is a driving force behind protein unfolding.[49-50] These chemicals are strong chaotropes, they disrupt the H-bonding network of the solvent water, disrupting the solvent-solvent interactions and in turn promoting solvent interactions with internal H-bonds in the protein core leading to protein unfolding. Chemical denaturants also interact with peptide bonds and sidechains of hydrophilic amino acid residues, thereby directly causing protein unfolding. Additionally, the application of two denaturing conditions simultaneously has compounding effects on protein unfolding. Extreme pH or the presence of denaturants greatly accelerate temperature induced protein unfolding as illustrated for model proteins Ribonuclease A and barnase.[49] Thus, solvent conditions are extremely important for maintaining the native state of proteins as elucidated by the important role of solvent-protein interactions.

The Anfinsen's dogma dictates that protein structure is decided wholly by the primary sequence[51] with there being two interesting exceptions: the class of intrinsically disordered proteins which are unfolded in physiological conditions[52] and the presence of chaperones/folding catalysts in the cell that affect the final fold of the protein (independent of the primary structure).[53] Protein folding in cells differs from folding *in vitro* in two major ways: a cytosolic environment which is extremely crowded, and the presence of molecular chaperones.[53] Chaperones bind to either nascently synthesized proteins or during later stages to aid protein folding in proceeding towards the native state in an ATP-dependent manner, thus allowing quality control of protein folding in the cell. Chaperones have an important role in protein transport and preventing the misfolding of proteins. The prevention of protein misfolding related proteinopathies is important for the cell as misfolding has been implicated in several neurodegenerative diseases.[47]

The misfolding of amyloidogenic proteins has been implicated in neurodegenerative diseases like Alzheimer's and Parkinson's disease. At least 20 proteins including the amyloid-β and α-synuclein have been shown to form toxic amyloid fibrils in humans. These fibrils are characterized by ordered filaments formed of cross β-sheets between the main chains of constituent polypeptides.[47] Soluble oligomers, which are the most toxic species are pre-fibrillar aggregates; they form protofilaments which come together and form the mature amyloid fibrils. In familial neurodegenerative diseases, mutations which increase hydrophobicity of amyloid forming protein or which makes the nascent protein more susceptible to proteolysis and truncation have been discovered. Thus in a cellular environment, a nascent polypeptide can have many diverse fates including correct folding, degradation due to misfolding and truncations, or the formation of toxic amyloid fibrils.

Protein unfolding is a prerequisite for targeting proteins towards ubiquitin-mediated proteasome degradation pathway. Another cellular process which requires protein unfolding is the translocation of proteins into the subcellular organelles such as mitochondria.[53] The involvement of protein unfolding in two key cellular processes and misfolding in pathogenic amyloid fibril formation thus illustrates the importance of understanding the phenomenon.

1.4.2 Techniques used to study protein folding

Transient ensembles observed in the early stages of protein folding are studied using a combination of fast triggering and detection methods.[21] One strategy to study protein folding involves the refolding of a denatured protein by transferring it from denaturing to non-denaturing solvent which will be the refolding trigger. Rapid mixing method (stopped-flow) to quickly change solvent conditions or non-mixing methods (Temperature jump, Electron transfer, and flash photolysis) can be used as the refolding trigger. The choice of refolding trigger depends on the timescale and must be faster than the biological process to be studied. Short pulse Lasers have been used to generate photochemical triggers and Temperature jumps. Photolysis of a CO-iron bond in cytochrome c and subsequent refolding was used to elucidate the importance of Met and His interactions in cytochrome c folding.[54] Additionally, IR Lasers have been used to effect Temperature jumps by exciting solvent molecules with a time resolution in the ns range.[55]

Horse Cytochrome c folding happens in the ms timescale. However, there is an initial rapid collapse of the polypeptide which occurs on timescales much faster (45—60 μs) than the rate-limited protein folding. Stopped flow CD was combined with intrinsic Trp fluorescence to study this burst-phase encountered in cytochrome c refolding. Significant resonance energy transfer from Trp to heme was observed, characterized by a 60 % decay in Trp fluorescence within the first 45—60 μs. This rapid burst phase was also elucidated for Ribonuclease A.[56]

Probing techniques provide the structural resolution in protein folding experiments. Protein folding involves a collapse of the polypeptide and packing of hydrophobic amino acid residues into the core (Figure 1.22). The absorbance and fluorescence of the intrinsic chromophores Trp, Tyr, and extrinsic chromophore 8-anilino-1-napthalensulphonate (ANS) are indirect probes for monitoring

polypeptide collapse. Trp burial in the protein core during folding accompanies a significant increase in fluorescence intensity and blue shift of emission maximum. The sensitivity of PhenylCyanine quantum yield to solvent allowed the use of this non-native aromatic amino acid analogue to study the insertion of a transmembrane peptide.[21]

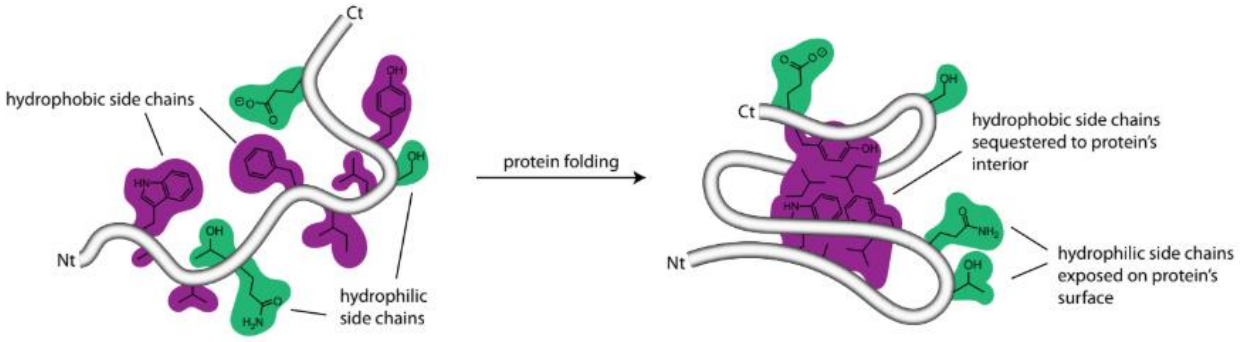


Figure 1.22: Hydrophobic collapse leading to the folding of a globular protein. Hydrophobic effect is one of the principle driving forces behind protein folding, leading to the formation of a core of hydrophobic amino acid residues and hydrophilic amino acid residues remaining solvent-exposed at the surface (Adapted from Dobson *et al*, 1996).[21]

In the presence of subdenaturing concentrations of GdnHCl, a molten globule-like state has been uncovered in the unfolding pathway of multiple proteins.[57] Ahmad *et al* illustrated the presence of this molten globule-like state for HSA at subdenaturing GdnHCl concentrations by monitoring ANS fluorescence increase in the presence of 0.5—2.5 M GdnHCl.[6] Native like secondary structure is expected to remain intact in molten globule states and the $MRE_{222\ nm}$ did not change in the presence of these denaturant concentrations.

Förster resonance energy transfer can be used to study folding thermodynamics and chain dynamics by monitoring the distance between a donor and an acceptor fluorophore. The choice of donor-acceptor fluorophore pair depends on the R_0 (Förster distance) and significant overlap between the donor emission and acceptor absorption spectra. Unnatural amino acids based FRET pairs have allowed the characterization of peptide dynamics. Fluorescence correlation spectroscopy (FCS) is based on the fluctuations in the fluorescence intensity of a fluorophore covalently attached to a protein molecule diffusing in and out of a very small observation volume. The correlations of these fluctuations happen either because of triplet state quenching or molecular diffusion. PET-FCS was used to study the chain dynamics of an Influenza A proton channel by monitoring Atto 655 quenching by surrounding Trp residues.[58]

Time-resolved fluorescence energy transfer was used to monitor intermediates in apomyoglobin refolding and thus allowing visualization of the heterogeneity of the protein folding pathway.[59] To retrieve exact information about molecular dimension of the transient conformational ensembles, other methods such as small-angle X-ray scattering (SAXS) and Dynamic Light Scattering (DLS) are used which give the molecule's Radius of Gyration.[21]

Linear IR spectroscopy uses vibrational transitions to deduce information about biomolecular dynamics happening in the ps timescales. The amide I band originates in the carbonyl stretching vibrations whose sensitivity to the protein fold can be used to characterize folding intermediates. Helix-turn-helix and beta-hairpin formation was studied using Temperature jump IR technique and the rate limiting step in the process was ascertained to be the turn formation.[58]

Numerous techniques exist to study the process of protein folding/unfolding. Significant differences may be observed in the characterization of protein unfolding by the conventional methods such as those discussed above. These arise due to differences in sensitivity of each technique to distinct aspects of molecular structure along with the heterogeneity of the protein unfolding pathways. Even though protein unfolding by conventional methods may reveal heterogeneous results, parameters obtained after analyzing protein unfolding by novel techniques need to be benchmarked against results obtained from the conventional methods.

ProCharTS absorbance and luminescence depends on the number of charged amino acid sidechains in close contact, whose number and nature are expected to decrease and change substantially in the protein unfolding process. Hence an attempt was made in this thesis to potentially use ProCharTS absorbance and luminescence for the LABEL-FREE detection of protein unfolding.

1.5 DNA-Protein binding

1.5.1 Sperm nuclear basic protein protamine

Protamines (PRM) are small, cationic peptides which replace histone as the DNA packaging material in maturing spermatids. They are characterized by an abundance of basic amino acid residues (50—70 %), presence of Arg clusters, and a remarkable sequence variability from species to species.[17] The most important function of PRM amongst other proposed functions is the compaction of the paternal genome, thus giving protection from nucleases and potential mutagens that the paternal genome may encounter while also making the spermatozoa more hydrodynamically compact allowing it to move faster towards the oocyte.[60]

Salmon protamine (PRM) is 32 amino acid residues long (marked with a red arrow in Figure 1.23) and contains 21 Arg including three Arg clusters. It is one of the well characterized protamines. An abundance of positively charged amino acid Arg (shown in blue), the absence of other charged amino acid residues and aromatic amino acid residues made it a suitable candidate for investigating the presence of recently reported novel Protein Charge Transfer absorbance and luminescence spectra. The percentage of charged amino acid residues in the protamines from multiple species is shown in red.

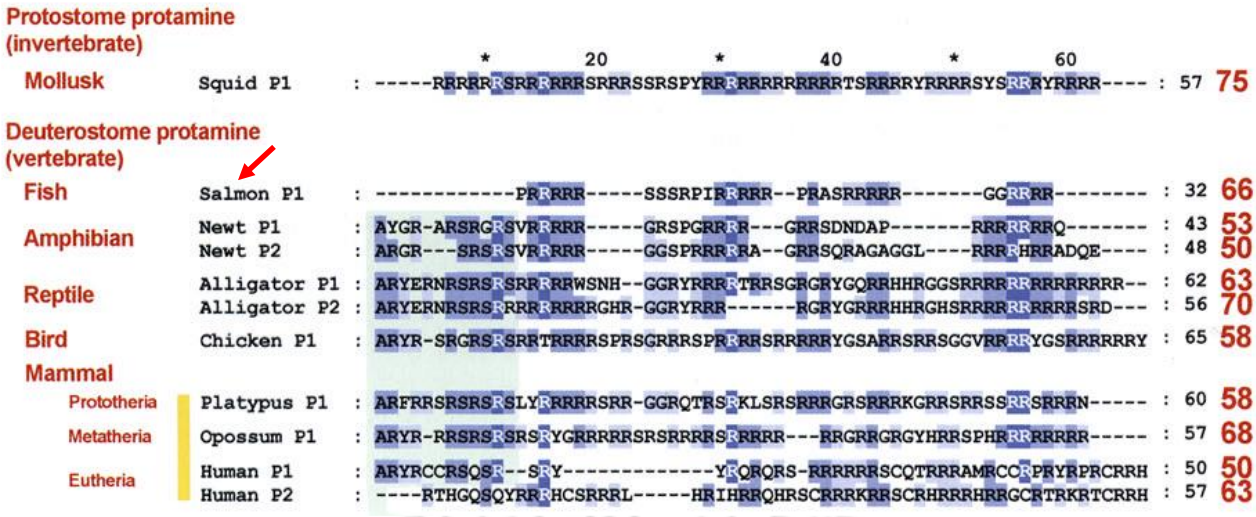


Figure 1.23: The amino acid sequence of protamine from different species. Protamines are characterized by an abundance of basic residues (shown in blue), absence of Trp residues, and significant variation in sequence between species (Adapted from Lewis *et al*, 2003).[61]

1.5.2 Salient features of the nucleoprotamine complex

PRM binding with DNA is hypothesized to happen via the Arg clusters binding to negatively charged phosphate backbone of DNA molecules. The B-form of DNA has been confirmed by Raman[62] and CD spectroscopy studies.[7] Based on vibrational bands in the Raman spectra of PRM-gDNA nucleoprotamine complex, the absence of A-form DNA was inferred.[62] B-form of DNA in nucleoprotamine complex was confirmed with CD by observing the characteristic positive band at ~275 nm and a negative band at ~245 nm in near UV CD spectra. Thus the secondary structure of DNA in nucleoprotamine complex is better understood than PRM secondary structure. Due to the absence of nucleoprotamine crystal structure, the structure and extent of PRM intramolecular H-bonding upon DNA binding has not been obtained conclusively.

The proposed nucleoprotamine complex structure is that one PRM molecule binds per helical turn of DNA at the major groove (Figure 1.24). PRM binds strongly to DNA, causing DNA condensation into toroidal subunits (Figure 1.24 b) which are 25 nm in thickness, have 50—70 nm diameter, and are condensed into structures that are even denser than nucleosomes containing 50,000 bp of DNA.[17] The absence of PRM structure in solution has been elucidated by many research groups however there is a lack of consensus on PRM structure upon DNA binding.

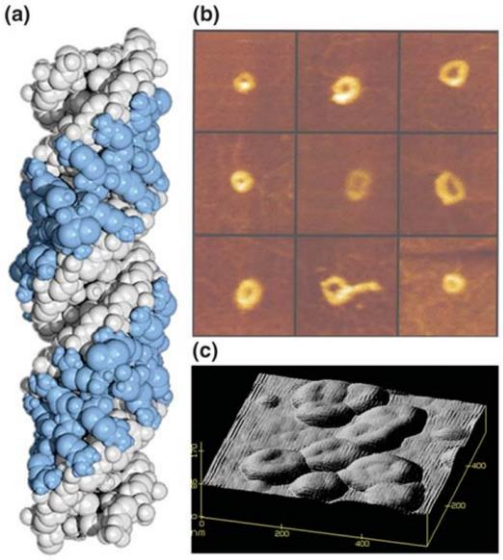


Figure 1.24: (a) Proposed model of nucleoprotamine structure showing one protamine molecule bound per turn of DNA helix. Scanning-probe microscopy images of (b) nucleoprotamine complexes prepared in vitro illustrating the toroidal shape and (c) complexes purified from sperm chromatin (Adapted from Balhorn, 2007).[17]

The two models proposed for PRM structure in nucleoprotamine complex are that 1) there is an absence of intramolecular H-bonding or 2) PRM adopts an α-helix conformation upon complex formation with DNA. The first model suggests that the guanidinium groups of Arg in PRM form electrostatic and hydrogen bonds with phosphate groups in DNA and there is an absence of PRM intramolecular H-bonding.[62]

According to the second model, PRM bound to DNA is rich in α-helix and turns. Based on FTIR studies done on PRM in salmon sperm chromatin, ~20 % of the PRM sequence was present in α-helix, 40 % in β-turns, and the remaining 40 % was found unstructured. PRM bound to chromatin in squid nucleus was also found to be structured with different secondary structure content reported.[63]

1.5.3 Techniques used to monitor DNA-Protein binding

Nucleic acid-protein binding is a crucial biomolecular interaction involved in regulation of transcription and other vital cellular processes. To understand DNA-protein binding various techniques are employed including EMSA, ITC, ChIP, and footprinting assays.

Electrophoretic mobility shift assays (EMSA) are widely used to qualitatively and quantitatively monitor DNA-protein binding. EMSA works on the principle that the electrophoretic mobility of protein-nucleic acid complexes will be lower than nucleic acid alone. Conventionally the protein of interest whose DNA binding needs to be investigated, is incubated with target DNA under standardized conditions and the complex formed is visualized after native acrylamide or agarose gel electrophoresis. Shifts in gel assays can be used to determine dissociation constants as well.[18]

The widespread application of EMSA to monitor DNA-protein binding results from the various advantages of the technique. EMSA assays can be performed with DNA and proteins of a wide range of sizes using very small sample amounts in various binding conditions. The use of P^{32} labelled DNA allows the sensitive detection of DNA-protein binding with very less protein and DNA concentrations (~1 nM and below).[64] Fluorescently labelled DNA is the method of choice when the working concentration of DNA is slightly higher. Alternatively, conventional detection by Ethidium Bromide (EtBr) can be used when DNA concentrations are greater than 5 ng per band. Thus various detection techniques are available to visualize nucleic acid after complex formation and electrophoresis making EMSA a method of choice for these applications.

EMSA is not without its drawbacks including 1) absence of information about nucleic acid site involved in protein binding, 2) labelling with fluorophores potentially influencing the nucleic acid-protein binding kinetics, 3) the possibility of dissociation of complex during electrophoresis, and 4) absence of exact correlation between shift observed and the size of the bound protein. The characterization of putative DNA binding sites is key to understanding how proteins such as transcription factors bind specifically to DNA. Since EMSA does not provide information about nucleotides involved in binding, footprinting methods are employed. DNase I footprinting assays are used to identify DNA-protein binding sites, employing the principle that DNA sequences bound to protein will be protected from the DNase I endonuclease activity.[64] After radioactive or fluorescent labelling of the target nucleotide sequence at one end, target DNA is cleaved and the fragments obtained after hydrolysis analysed on denaturing polyacrylamide gels. The fragments corresponding to binding site will be absent as they are protected from DNase I hydrolysis compared to the control in which DNA binding protein was not added. Thus footprinting assays will reveal the targets of site-specific DNA binding proteins.

Surface plasmon resonance (SPR) is used to monitor biomolecular interactions and has been applied to study DNA-protein binding.[65] It allows the real time monitoring of kinetics and binding affinity of DNA for analyte protein(s). This label-free technique involves the immobilization of DNA molecules on a chip surface coated with a metal (such as gold which is conducive for surface plasmon resonance) and then allowing binding with protein to obtain a sensogram. A sensogram consists of the association, equilibrium, dissociation, and regeneration phase from which the dissociation constant can be determined using the association and dissociation rates. SPR can also be used to monitor stoichiometry of the binding system. SPR works on the principle of total internal reflection and is extremely sensitive to the refractive index changes at metal surface. The refractive index increases proportionately to the mass increase upon analyte-DNA binding which results in an increase in the resonance angle (at the resonance angle, there is maximum resonance of surface plasmons with incident light resulting in lowest intensity of reflected light). This increase is reported in response units and is maximum when all binding sites on DNA are occupied. SPR has been used extensively to characterize DNA-protein binding and very effectively to study multiprotein-DNA complexes.

Chromatin Immunoprecipitation (ChIP) is used to determine physiological protein-gene interactions and post-translational modification status of associated chromatin. The procedure involves chemical crosslinking of protein and associated chromatin in living cells to preserve relevant protein-chromatin interactions. After crosslinking, chromatin shearing is done and smaller fragments obtained are probed using antibodies specific for the protein of interest.[64] The DNA-protein complex is immunoprecipitated, sample crosslinks reversed and the presence of specific DNA fragments determined after PCR. The binding sites of transcription factors of many species have been determined in vivo using ChIP thus illustrating the scope of the technique.[66]

Isothermal titration calorimetry (ITC) allows the detailed thermodynamic characterization of interaction between biomolecules. The binding of DNA and protein is accompanied by bond rearrangements and the resulting change in Gibbs free energy from the interaction is measured by ITC. To perform ITC experiment, ligand (protein in the case of DNA-protein interactions) is injected into the sample cell containing DNA. The heat energy per unit time required to keep the sample cell at thermal equilibrium with reference cell (containing buffer) after each injection of ligand is measured.[67] Upon injection of ligand, if the reaction is exothermic, then less heat energy is required to maintain the temperature and the peaks will be negative as shown in Figure 1.25. The peaks will be positive in the case of an endothermic reaction. Aliquots of ligand are added to the sample until the heat input goes to the baseline. The area under each peak will give the heat change for that injection. The observed heat change per injection can be used to obtain thermodynamic parameters for a DNA-protein binding experiment by fitting to the theoretical binding curve. Thermodynamic parameters enthalpy, binding constant, and stoichiometry can be obtained by nonlinear regression analysis of the theoretical binding curve.

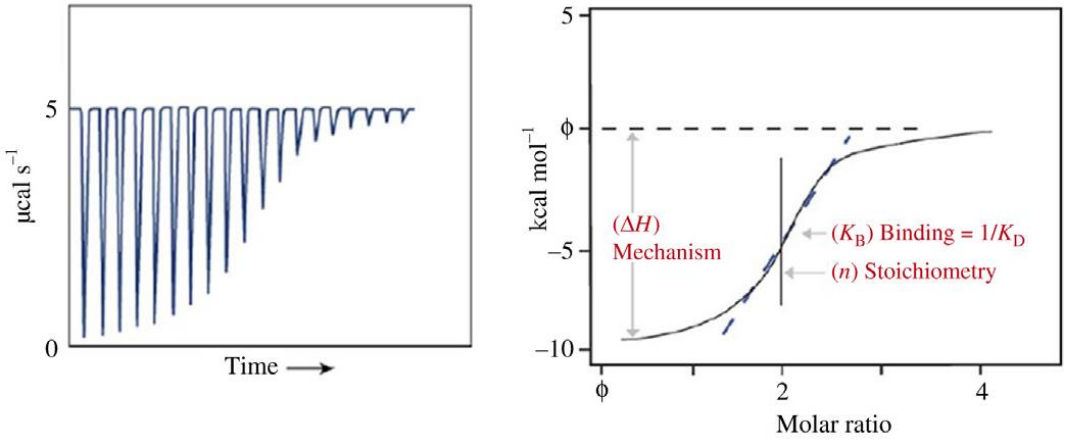


Figure 1.25: A typical ITC thermogram showing the DNA-protein binding parameters that can be obtained from an ITC experiment (Adapted from Srivastava *et al*, 2019).[68]

The circular dichroism of DNA from 200—300 nm has been hypothesized to originate from three sources: the chirality of the monomeric nucleosides owing to the asymmetric sugars, the chirality of nucleic acids upon helical secondary structure formation, and the formation of long-range tertiary interactions.[69] In physiological conditions, DNA is most abundantly present in the B-form. In B-form DNA, the composite base pairs are aligned perpendicular to the DNA axis resulting in

weak chirality. The characteristic CD spectrum of B-form DNA has a weak but prominent positive band at ~275 nm and a negative band at ~245 nm with similar intensity to the ~275 nm band.

The disordered nature of strands in denatured DNA implies low chirality and significantly reduced circular dichroism which has been used to characterize B-form DNA helix-to-coil transition upon thermal denaturation. The distinctive S-shaped curve obtained when helix-to-coil transition was monitored by following CD at a fixed wavelength has been used to determine melting point (T_m). The B, A, and Z-forms of DNA have been identified to be biologically active with the A-form DNA being present in dehydrating conditions. Because of an inclination of the base pairs to helix axis, A-form DNA is more chiral compared to B-form with a strong positive band at ~260 nm and a strong negative band at ~210 nm. The B—A transition induced in the presence of increasing amounts of ethanol (such as trifluoroethanol) has been characterized by monitoring CD spectra.[69] Additionally B—Z transition, hairpin formation, guanine, and cytosine quadruplex formation has been demonstrated using CD spectroscopy. These transitions are followed by monitoring the disappearance of peaks characteristic of B-form DNA concomitant with the appearance of peaks in the 200—300 nm region which are characteristic features of the alternate DNA conformation.

The change in CD at 275 nm upon nucleoprotamine complex formation was used to monitor salmon PRM binding with calf thymus DNA. Upon incubation of calf thymus DNA with increasing amounts of PRM, a reduction in intensity at 275 nm and a red shift of the positive band maximum and crossover wavelength was observed.[7] Similar CD changes have also been observed upon DNA binding with polylysine and in chromatin. Thus CD has demonstrably been used to monitor structural transitions in nucleic acids and these studies were the basis for potentially employing CD spectroscopy to monitor the formation and characterization of nucleoprotamine complex in this study as well.

Protamine binds non-specifically to DNA and causes DNA condensation. Alternatively, the inability of BSA to condense DNA has been shown.[20] DNA condensation assays monitor the decrease in DNA absorbance after condensation and subsequent precipitation by PRM binding.[19] The assay relies on the observation that DNA condensation by PRM inherently leads to precipitation of the nucleoprotamine complex. The DNA condensation process cannot be rigorously distinguished from the nucleoprotamine complex precipitation.[70] The assay is performed by keeping the DNA concentration constant and mixing increasing amounts of PRM. After DNA-protein mixing, the samples are centrifuged, and absorbance of supernatant is monitored (usually at 260 nm which is approximately B-DNA absorbance maximum). The condensation of genomic DNA by salmon protamine was illustrated by Gupta *et al* upon monitoring the decrease in DNA absorbance of the gDNA-PRM mixture supernatants.[19] At higher PRM concentrations, nucleoprotamine complex formation and subsequent precipitation increased which reduced the effective DNA concentration in supernatant, allowing the binding of DNA-protein to be monitored.

Since clarity on PRM structure in nucleoprotamine complex is still elusive and because ProCharTS depends on the number of charged Arginine sidechains in close contact, it is worthwhile to investigate whether PRM ProCharTS can be used to detect structure gain upon DNA-protein binding. The possibility of PRM ProCharTS decreasing upon DNA binding due to neutralization of charges in PRM could additionally be used to detect DNA binding.

1.6 Objectives of thesis work

Multiple reports of novel absorbance and luminescence in non-aromatic amino acids and proteins were discussed in the literature review. The observation of novel intrinsic protein absorbance and luminescence by multiple research groups prompted an investigation into the presence of ProCharTS in charge rich proteins α_3C, α_3W, and PRM. Before understanding potential applications of ProCharTS, the effects of solvent parameters such as pH and salt concentration on the spectra needed to be characterized. Additionally, due to the overlap between absorbance and luminescence of CT states with Trp absorbance and fluorescence, the possible effects of ProCharTS luminescence on Trp fluorescence was monitored in α_3W. After sufficient understanding of the effect of solvent parameters on ProCharTS was obtained, potential applications of ProCharTS to monitor protein unfolding and DNA binding were investigated. The applications of intrinsic Protein Charge Transfer Spectra to monitor protein unfolding and DNA binding in a LABEL-FREE method were elucidated.

The objectives of thesis work were thus decided as follows:

1. **Observing Protein Charge Transfer Spectra in proteins α_3C, α_3W, and PRM.**
2. **Monitoring the effect of solvent parameters on ProCharTS and investigating the influence of ProCharTS on Trp fluorescence in proteins.**
3. **The application of Protein Charge Transfer Spectra to monitor protein unfolding.**
4. **Monitoring DNA-protein binding using Protein Charge Transfer Spectra.**

1.7 Thesis at a glance

The thesis work was titled "**Monitoring Protein Unfolding and DNA Binding using Protein Charge Transfer Spectra"** and was carried out to achieve four major objectives. The major sections of the thesis including review of literature, experimental methods, and results of the work performed are shown in the underlying Figure 1.26.

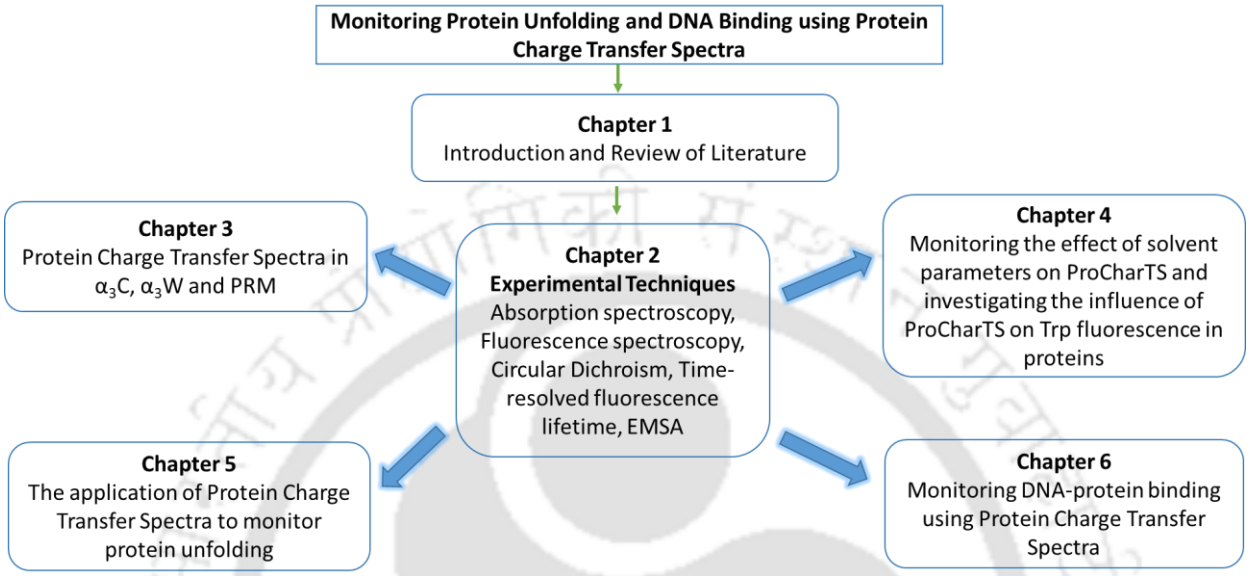


Figure 1.26: Thesis at a glance

CHAPTER 2

EXPERIMENTAL TECHNIQUES, MATERIALS, AND METHODS

2.1 Experimental techniques

The principles behind the spectroscopic techniques including absorbance, steady-state fluorescence, time-resolved fluorescence, and circular dichroism spectroscopy used in the thesis work are described below.

2.1.1 UV-Visible spectroscopy

Spectroscopy deals with the measurement and analysis of how electromagnetic radiation interacts with matter. Light consists of oscillating electromagnetic fields which upon interaction with matter can promote electrons in the ground state to excited states. The transition from ground to excited states in a chromophore happens when the energy of the incident radiation is equal to the energy difference between the two states. Energy of the incident radiation is given by:

$$E = \frac{hc}{\lambda} = h\nu \qquad 2.1$$

where h = Planck's constant, c = speed, λ = wavelength, and ν = frequency of incident radiation. The molecule in the excited state can relax to the ground state either non-radiatively dissipating the energy as heat or via the emission of a lower energy photon through fluorescence/phosphorescence. Light can also be scattered (Raman scattering) or reflected by matter. UV-Visible absorbance and fluorescence spectroscopy is routinely used to characterize protein structure and function along with other techniques such as nuclear magnetic resonance (NMR), Fourier-transform infrared (FTIR), and circular dichroism (CD) spectroscopy.

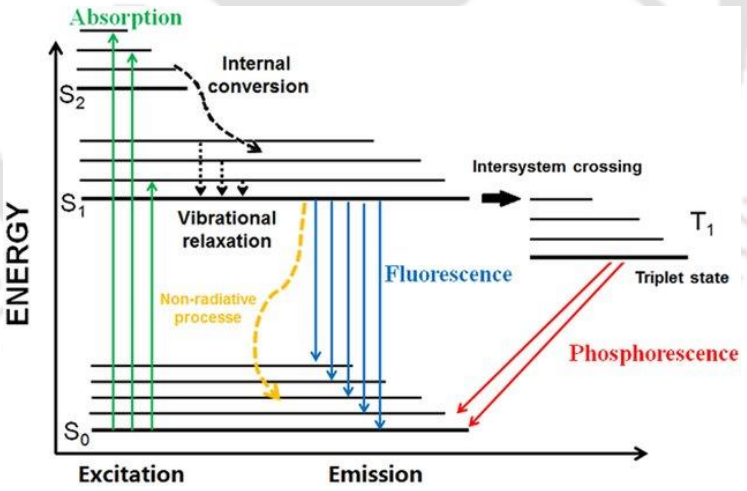


Figure 2.1: Jablonski diagram depicting the absorption and emission of light (Adapted from Schweizer *et al*, 2021).[71]

The absorbance and emission of UV-Visible light is described by a Jablonski diagram. Jablonski diagram illustrating the electronic states and the associated vibrational states of a representative fluorophore is shown in Figure 2.1. The molecular processes involved in the absorption and emission of photon are described. The absorption of photon happens in the fs timescale and occurs according to the Frank-Condon principle (absorption happens without nuclear reorganization). At room temperature, absorption and emission happens from the lowest vibrational state. Upon

absorption of photon, the electronic transition will happen from the lowest vibrational state of S_0 to the vibrational states of S_1 or S_2. Fluorescence or phosphorescence will happen from S_1 or T_1 with the emission of a photon having lower energy than incident radiation. Internal conversion, intersystem crossing, vibrational relaxation, and fluorescence quenching are the non-radiative processes.

2.1.1.1 Absorbance spectroscopy

The interaction of light with matter can cause electronic excitation from the ground state to higher electronic states. The process occurs in fs timescale and depends on the energy gap between electronic states of the molecule and the wavelength of incident radiation. The intensity of light absorbed is proportional to the molecular parameter transition dipole moment. The transition dipole moment is $< \psi_b|\tilde{\mu}|\psi_a >$.

ψ_a and ψ_b denote the ground and excited state wavefunctions of a chromophore. Upon interaction with light, a fraction of molecules transition from the ground state to the excited state. The electric field of the incident radiation perturbs the electric charge distribution in the chromophore. The rate at which this electronic transition happens is an important parameter and is calculated using:

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

where B_{ab} = transition rate per unit energy density and I(ν) = energy density incident on molecule. The Einstein coefficient B_{ab} is related to the transition dipole moment by the following equation:

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

where ħ = reduced Planck's constant. The molar extinction coefficient can be related to molecular properties using the equation:

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

where c = velocity of light, h = Planck's constant, and N_0 = Avogadro's number.

Using equation 2.5, dipole strength D_{ab} can be calculated from B_{ab}.

$$| < \psi_b|\tilde{\mu}|\psi_a > |^2 = 9.18 * 10^{-3} \int \left(\frac{\varepsilon'}{\nu}\right) d\nu = D_{ab} \qquad 2.5$$

Oscillator strength compares the absorbance intensity of a chromophore to that expected from a three-dimensional harmonic oscillator. Oscillator strength gives the probability of transition from ground state to excited state and is in the range of 0.1 to 1 for strongly allowed transitions and can be calculated using equation 2.6.

$$f_{ab} = (8\pi^2 mc/3h\nu) D_{ab} = 4.315 * 10^{-9} \int (\varepsilon'/\nu)\, d\nu \qquad 2.6$$

The fraction of light absorbed is proportional to the number of absorbing molecules and this relationship is described by the Beer-Lambert's law. Extinction coefficient (ε) is a measure of how strongly a chromophore absorbs at a particular wavelength and is independent of concentration.

The intrinsic chromophore Trp absorbs more strongly than Tyr with Trp ε_{max} = 5600 $M^{-1}cm^{-1}$ and Tyr ε_{max} = 1400 $M^{-1}cm^{-1}$.

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

where I_0 and I are the incident and transmitted light intensities. The absorbance of chromophores in the UV-Visible region is monitored using spectrophotometers. The light source, optical components, detector, sample and reference cell of a typical dual-beam UV-Visible spectrophotometer are illustrated in Figure 2.2.

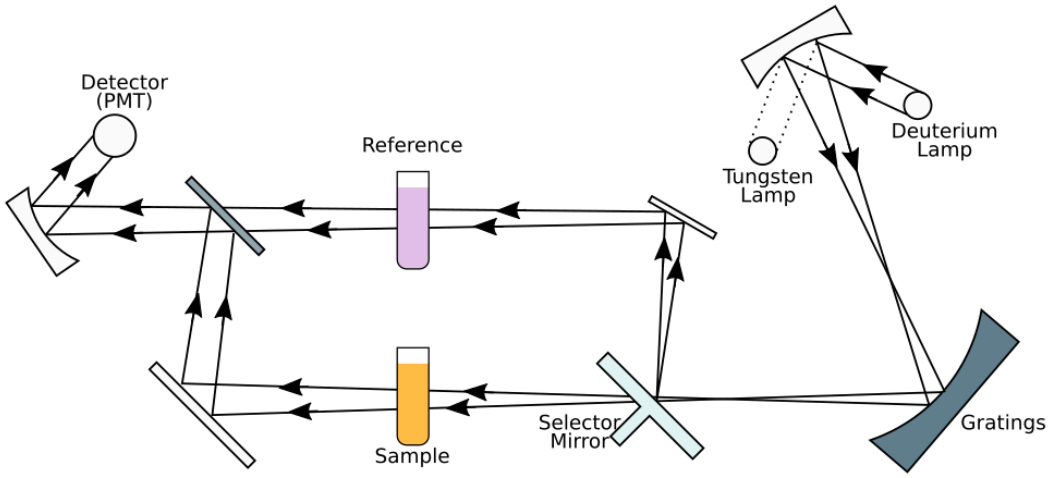


Figure 2.2: Schematic representation of a dual-beam UV-Visible spectrophotometer (Adapted from Biophysical Chemistry II, 1980).[8]

2.1.1.2 Fluorescence spectroscopy

Fluorescence is the relaxation process which involves emission of UV or visible light after absorption of light by fluorophore. Fluorescence represents the transition of electron from the first excited singlet (S_1) to ground singlet state (S_0) accompanying the emission of a photon with lower energy than the incident radiation. The emission of photon happens relatively rapidly (in the ns timescale) and in proteins involves the aromatic amino acid residues especially Trp. Compared to absorbance (fs timescale), fluorescence happens at a slower timescale (ns) and hence is more sensitive to processes such as solvent exposure, protein rotational diffusion, and quencher diffusion.

2.1.1.2.1 Internal conversion and intersystem crossing

Upon absorbing light, a fluorophore typically gets excited to a higher vibrational level of S_1 or S_2. Prior to emission, chromophore relaxes to lowest vibrational state of S_1 in the ps timescale by a process known as internal conversion. Internal conversion happens faster than fluorescence and all the molecules are present in S_1 lowest vibrational state at the time of emission. The spin conversion of molecules from S_1 to first triplet state T_1 occurs by intersystem crossing. Compared to fluorescence, the electronic transition from T_1 to ground state is shifted to longer wavelengths. The phenomenon is called phosphorescence and due to the spin forbidden nature of the transition,

rate constants for phosphorescence are much lower than fluorescence. The processes of internal conversion and intersystem crossing are illustrated in Figure 2.1.

Thus non radiative decay includes internal conversion, intersystem crossing, and fluorescence quenching: $k_{nr} = k_{isc} + k_{ic} + k_q[Q]$; where k_{nr}, k_{isc}, k_{ic}, k_q are the rate constants of non-radiative decay, intersystem crossing, internal conversion, quenching and [Q] is the quencher concentration.

The observed fluorescence lifetime or the average time a fluorophore spends in excited state is defined as:

$$< \tau_f >= \frac{1}{k_r + k_{nr}} \qquad 2.8$$

where k_r and k_{nr} are rate constants of radiative and non-radiative decays. Quantum yield measures the photon emission efficiency of a fluorophore and is the fraction of absorbed photons emitted. Quantum yield of a fluorophore would hence depend on the rate constants of radiative and non-radiative decays:

$$\phi_f = k_r/(k_r + k_{nr}) = k_r/(k_r + k_{isc} + k_{ic} + k_q[Q]) \qquad 2.9$$

Quantum yield is related to the fluorescence lifetime by the following relationship:

$$\phi_f = \frac{\tau_f}{\tau_n} \qquad 2.10$$

where τ_n is the natural lifetime of a fluorophore (lifetime in the absence of non-radiative decay processes).

2.1.1.2.2 Characteristics of fluorescence

Two important photophysical characteristics of fluorescence are the presence of Stokes shift and the emission spectra being independent of excitation wavelength. The emission from a fluorophore happens at lower energy and longer wavelength compared to absorbance. This shift in emission maximum compared to absorbance maximum is called Stokes shift and is attributed to internal conversion and solvent relaxation.

Kasha's rule states that shape of the fluorescence emission spectrum is independent of the excitation wavelength used. The independence of emission maximum from excitation wavelength occurs because emission happens only from the lowest vibrational state of S_1. Upon excitation of molecules from the ground state to one of the vibrational states of S_1 or S_2, the excess energy is lost by internal conversion. It is interesting to investigate whether the novel ProCharTS luminescence shows these two characteristic features of fluorescence. Differences and similarities compared to fluorescence will shed light on the nature of the electronic states involved in observed novel luminescence.

2.1.1.2.3 Steady-state fluorescence

Steady-state fluorescence measurements require constant illumination of the sample followed by recording the fluorescence intensity at emissive wavelengths to collect the emission spectra. Due to constant illumination, during steady-state measurements fluorophore populations in the ground and excited state are in equilibrium. Steady-state measurements of anisotropy, fluorescence quenching, and resonance energy transfer provide information about protein dynamics. The sensitivity of a fluorophore to its environment forms the basis of its application to monitor protein structure and function. The change in polarity upon protein unfolding and exposure of buried Trp to solvent is accompanied with a red shift in emission maximum and reduction in fluorescence intensity. The steady-state fluorescence is the average of the time-resolved fluorescence and is directly proportional to the fluorophore lifetime.

$$I_{ss} = \int_0^\infty I_0 e^{-t/\tau} dt = I_0 \tau \qquad 2.11$$

where I_0 = intensity at time 0 and τ = fluorophore lifetime.

2.1.1.2.4 Steady-state anisotropy

Steady-state anisotropy is used to monitor the rotational mobility of fluorophores. The rotational diffusion of intrinsic chromophore Trp or extrinsic chromophores conjugated to sidechains of amino acid residues in proteins give information about the shape and size of the proteins. Changes in steady-state anisotropy can be used to monitor protein folding, protein-protein interactions, protein aggregation, and the partitioning of fluorophores into lipid bilayers. Fluorophores are illuminated with vertically polarized light and the extent of depolarization in the emitted light reveals the rotational mobility of the fluorophore.

The steady-state anisotropy is related to the τ and θ (rotational correlation time) of a fluorophore by the Perrin's equation:

$$r = \frac{r_0}{1+\frac{\tau}{\theta}} \qquad 2.12$$

where r_0 is the fundamental anisotropy of the fluorophore, obtained when rotational diffusion is absent. Rotational diffusion reduces the fluorophore anisotropy. The rotational correlation time of fluorophore in non-viscous solutions is low compared to the fluorophore lifetime. This implies that fluorophore would have rotated multiple times due to Brownian motion in the excited state and the emission obtained would be depolarized. The diffusion of fluorophore is greatly reduced upon attachment to proteins and the anisotropy increases with protein size. The rotational correlation time is sensitive to the local and global rotational motion of proteins.

Anisotropy measurements are made after the photo-selective excitation of fluorophores using polarized light. Figure 2.3 illustrates that only those fluorophores will absorb photons whose transition moment is aligned parallel to polarized light. In the absence of energy transfer, rotational

diffusion will determine how depolarized the emission and how low the anisotropy value would be. Steady-state anisotropy is related to the polarized emissions by the following equation:

$$r_{ss} = \frac{I_{||} - GI_{\perp}}{I_{||} + 2GI_{\perp}} \qquad 2.13$$

Here r_{ss} denotes steady-state anisotropy, $I_{||}$ and $I_{\perp}$ are the fluorescence intensity when the emission polarizer is parallel and perpendicular to the excitation polarizer. G = $\frac{S_V}{S_H}$ is the G-factor and corrects for the different sensitivity of detector and monochromator to vertically (S_V) and horizontally (S_H) polarized light.

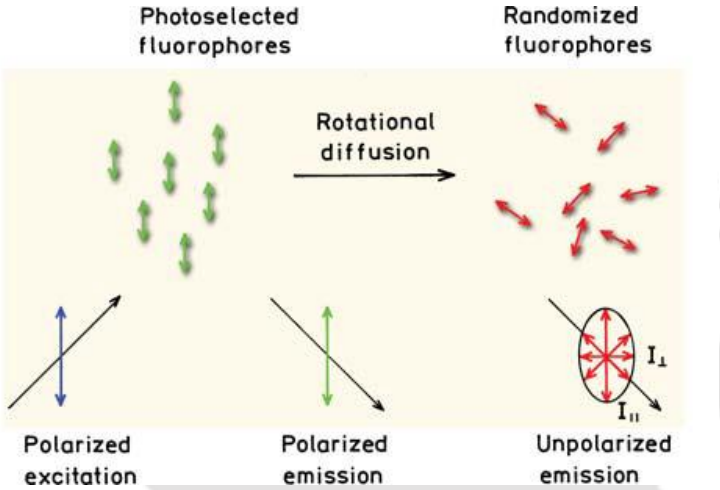


Figure 2.3: Rotational diffusion of fluorophores can be characterized by monitoring steady-state anisotropy (Adapted from Principles of Fluorescence Spectroscopy, 2004).[23]

2.1.1.2.5 Time-resolved fluorescence intensity decay

Time-resolved fluorescence measurements are used to extract lifetime and rotation correlation time parameters from fluorescence intensity decays. Two important applications of fluorescence lifetime measurements include studying the dynamics of biological macromolecules and fluorescence lifetime imaging microscopy. Time-domain fluorescence measurements are made using time-correlated single photon counting (TCSPC) technique. TCSPC employs a sub-nanosecond or picosecond pulsed excitation source with pulse width much shorter than fluorophore lifetime to excite samples. For the obtained fluorescence decay, the slope of log intensity plotted against time is used to obtain τ of the excited fluorophore. To avoid the influence of rotational diffusion and anisotropy, fluorescence intensity decays are measured using a polarizer at 54.7 ° angle as shown in Figure 2.4.

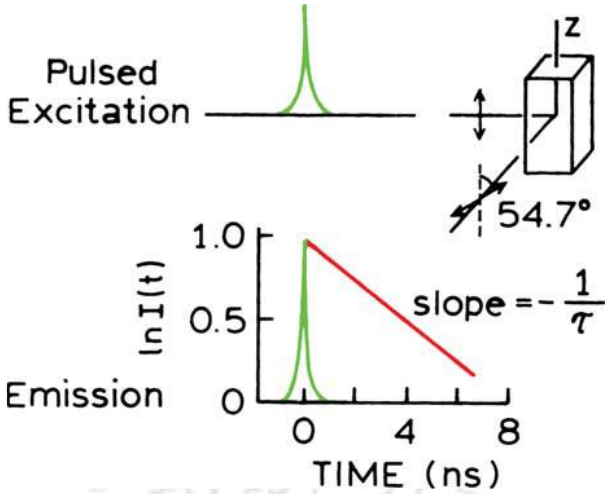


Figure 2.4: Schematic representation of time-domain fluorescence measurements (Adapted from Principles Of Fluorescence Spectroscopy, 2004).[23]

The lifetime of a fluorophore (the average time spent in excited state) can be calculated by averaging the time <t> using the following relationship:

$$< t > = \frac{\int_0^\infty tI(t)dt}{\int_0^\infty I(t)dt} = \frac{\int_0^\infty t\, e^{-t/\tau}dt}{\int_0^\infty e^{-t/\tau}dt} \qquad 2.14$$

For a single exponential decay, the τ = <t> and the intensity decay is related to the fluorescence lifetime by the following relationship: $I = I_0 e^{-t/\tau}$. 2.15

When two fluorophores are present in the biomolecule or a fluorophore is present in a heterogeneous environment, the fluorescence intensity decay becomes multi-exponential. The intensity decay of a multi-exponential fluorophore is related to individual lifetimes by the following relationship:

$$I(t) = \sum_i \alpha_i e^{-t/\tau_i} \qquad 2.16$$

The mean lifetime of a multi-exponential decay is $\tau_m = \sum_{i=1}^{n} \alpha_i \tau_i$ where i is generally 2 or 3 but can be higher with the constraint that sum of amplitudes is 1.

Advantages of time-resolved over steady-state fluorescence measurements are that: (1) fractional fluorescence contribution from each fluorophore when two fluorophores are present can be ascertained, (2) individual components can reveal the multiple conformations a biomolecule is present in and (3) lifetime measurements are independent of concentration. However, the fluorescence intensity decay analysis is not straightforward and is discussed in the following section.

2.1.1.2.6 Time-resolved intensity decay analysis

The fluorescence intensity decay analysis is associated with three curves: (1) measured data $N(t_k)$, the IRF (instrument response function) $L(t_k)$, and the calculated decay $N_c(t_k)$. The obtained functions are in discrete times (t_k) because in a TCSPC setup, photons are counted in channels with known width (Δt) and time (t_k). The IRF is collected using a dilute scattering solution (for

example, a suspension of chalk powder) and instrument response to this sample with zero lifetime is measured.[23] The IRF represents the shortest lifetime that can be resolved in that particular TCSPC instrument.

The measured intensity decay $N(t_k)$ is convoluted with the IRF and to obtain lifetime and amplitude parameters, intensity decay $N_c(t_k)$ needs to be calculated. The intensity decay $N_c(t_k)$ is calculated using the mathematical concept of iterative reconvolution after measuring the intensity decay $N(t_k)$ and IRF. Iterative reconvolution starts with the assumption that the excitation pulse is a series of δ-functions with varying amplitudes. The impulse response from sample at time t_k for each excitation δ-function is:

$$I_k(t) = L(t_k)I(t - t_k)\Delta t \qquad 2.17$$

where $L(t_k)$ is the excitation intensity, Δt is the channel width ($t > t_k$) in which photon counting is done and $I(t\text{-}t_k)$ measures intensity from time t to t_k. The measured decay $N(t_k)$ is the sum of the impulse responses created by all the individual δ-function excitation pulses occurring until t_k.[23]

The impulse responses obtained by the individual excitation pulses are summed up to obtain the measured decay $N(t_k)$:

$$N(t_k) = \sum_{t=0}^{t=t_k} L(t_k)I(t - t_k)\Delta t \qquad 2.18$$

When the channel width is small, the measured decay takes form of the equation:

$$N(t_k) = \int_0^t L(t')I(t - t')dt' \qquad 2.19$$

This relationship implies that the measured intensity is the sum of intensities obtained from all the δ-function excitation pulses.

The lifetime of a single exponential decay is the τ value for which the calculated time-dependent intensities $N_c(t_k)$ best fit the measured data $N(t_k)$. In multi-exponential decay analysis, the objective is to obtain α_i and τ_i values for which calculated intensities best fit the measured data. Nonlinear least square (NLLS) and maximum entropy method (MEM) are used to analyse fluorescence intensity decays.

Fit parameters are obtained from measured intensity by applying Marquardt's method to NLLS analysis.[72] For a multi-exponential decay, the first step is to use τ_i and α_i guess values to calculate initial intensity decay. Then through the use of iterative reconvolution, τ_i and α_i values are corrected such that $N_c(t_k)$ fits well to measured $N(t_k)$. The goodness of fit is analysed by the fit parameter χ^2:

$$\chi^2 = \sum_{k=1}^{n} \frac{1}{\sigma_k^2}[N(t_k) - N_c(t_k)]^2 = \sum_{k=1}^{n} \frac{[N(t_k) - N_c(t_k)]^2}{N(t_k)} \qquad 2.20$$

Since χ^2 is dependent on the number of data points, reduced χ^2 is a better parameter to assess the quality of fitted data:

$$\chi_R^2 = \frac{\chi}{n-p} = \frac{\chi^2}{\nu} \quad 2.21$$

where n = number of data points and p = degrees of freedom.[23] When the calculated data fits well to the observed decay, reduced χ^2 is between 1—1.2. Apart from reduced χ^2, residual distribution is visually examined to assess goodness of fit.

The difference between measured and fitted data is reflected in the standard deviation at each timepoint. The deviation is calculated using the formula 2.22 with time and an observed randomness in residual distribution implies that fit parameters obtained are close to correct values and the assumed model for data fitting is correct.

$$D_k = \frac{I(t_k) - I_c(t_k)}{\sqrt{I(t_k)}} \quad 2.22$$

Discrete analysis assumes that the intensity decay fits to a single, sum of two or three exponentials model and obtains the fit parameters.

Acquiring information about lifetime distribution requires that assumptions about the model are not made. Maximum entropy method (MEM) analysis is used to analyse and retrieve lifetime distributions.[73] In MEM fitting, the fluorescence lifetimes are assumed to be uniformly distributed in the 0.02—10 ns range at the beginning of the analysis. In each iteration of data fitting, the χ^2 is attempted to be minimized and Shannon-Jaynes entropy[36] (Formula 2.23) maximized:

$$S = \sum_i p_i log p_i \quad 2.23$$

where p_i denotes probability value of the i^{th} lifetime component. The obtained lifetime distributions are deemed to be correct when χ^2 is between 1—1.2, entropy is high and residuals are distributed randomly. The χ^2 for MEM fitting[36] is calculated using:

$$\chi^2 = \frac{(1/m)\sum_{i=1}^{m}[F_c(t_i) - F_e(t_i)]^2}{F_e(t_i)} \quad 2.24$$

2.1.2 Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy is used to determine the secondary structure content of proteins and nucleic acids. Peptide bonds present in an achiral environment such as an alpha-helix show circular dichroism, which is the unequal absorption of right and left handed circularly polarized light. Ellipticity (θ) is calculated form circular dichroism using the formula:

$$\theta = 2.303(A_L - A_R)180/4\pi \; degrees \quad 2.25$$

Rotational strength of a chromophore can be computed from molecular properties and is a measure of how strongly dichroic the chromophore is. The rotational strength is analogous to the dipole strength parameter of absorbance.[8]

$$R_{0a} = (3hc/8\pi^3 N_0) \int \{[\theta(\lambda)]/\lambda\} d\lambda \quad 2.26$$

R_{0a} = rotational strength of ground to excited state transition, h = Planck's constant, c = velocity of light, and N_0 = Avogadro's number.

Upon interaction with circularly polarized light, a chiral molecule will transition from ground state ψ_0 to excited state ψ_a and the rotational strength is calculated using:

$$R_{0a} = imaginary(<\psi_0|\tilde{\mu}|\psi_a>.<\psi_a|\tilde{m}|\psi_0>) \quad 2.27$$

where $\tilde{\mu}$ and $\tilde{m}$ are the electric and magnetic dipole moment operator.

The ellipticity of a chromophore is converted to molar or mean residual ellipticity (molar or residue basis) units which are independent of concentration. Molar ellipticity is related to the extinction coefficient by the following relationship:

$$[\theta] = 3300\ \Delta\varepsilon \quad 2.28$$

where $\Delta\varepsilon$ represents the change in absorbance ΔA corrected for protein concentration.

Peptide bonds are the chromophore used to perform far UV CD spectroscopy to determine protein secondary structure. In the 190—240 nm region, secondary structures such as alpha-helix, beta-sheet or random coil produce characteristic features in the CD spectra (Figure 2.5). The peptide bond absorbs strongly in the ~190 nm region corresponding to higher energy π to π * transition and a weak n to π * transition in the 210—220 nm region.

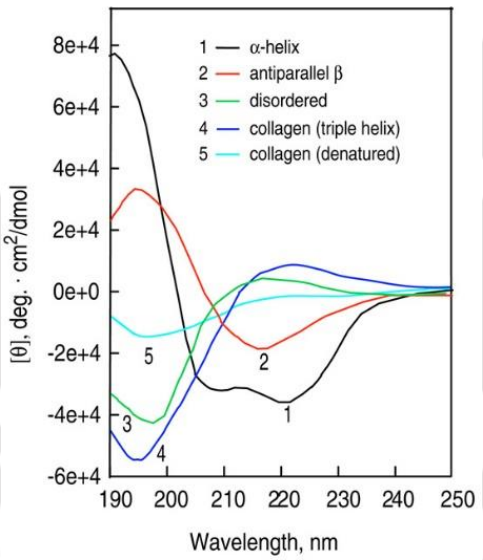


Figure 2.5: Secondary structure elements in protein and their characteristic features in the CD spectrum (Adapted from Greenfield, 2006).[74]

Near UV CD spectra is used to study the tertiary structure of proteins. The aromatic amino acids Trp, Tyr, and Phe contribute to the circular dichroism in the near UV region (250—300 nm). In a folded protein, aromatic amino acid sidechains can be placed in chiral environments which are monitored using near UV CD.

2.2 Materials

Acrylamide (A3553), ammonium persulfate (A3678), N,N,N′,N′-Tetramethylethane-1,2-diamine (A7024), gel loading buffer (G2525), β-mercaptoethanol (63689), agarose (A9539), EtBr (E7637), ampicillin (A8351), isopropyl β-D-1-thiogalactopyranoside (I5502), phenylmethylsulfonyl

fluoride (P7626), Calcium chloride (C8106), glycine (50046), sinapic acid (85429), trifluoroacetic acid (T6508), acetonitrile (A3553), guanidine hydrochloride (G3272), 9,10-diphenylanthracene (D205001), and N-acetyl L-tryptophanamide (A6501) were purchased from Sigma Aldrich. The proteins bovine serum albumin (A3059), human serum albumin (A1653), hen egg white lysozyme (L6876), protamine (P4005), and DNA sodium salt from calf thymus (D1501) were also procured from Sigma.

IAEDANS (5-({2-[(iodoacetyl)amino]ethyl}amino)naphthalene-1-sulfonic acid) (I14) and Dansyl (5-(DimethylAmino)Naphthalene-1-SulfonYL) Chloride (D73) were purchased from Invitrogen. Luria agar (M557), terrific growth medium (G004), Luria broth (M1245), magnesium chloride (TC186), and imidazole (GRM1864) were purchased from Himedia laboratories. PD-10 desalting column (17-0851-01) was procured from GE healthcare. Sodium acetate (17952), tris buffer (93315), citric acid monohydrate (100244), sodium dihydrogen phosphate monohydrate (106346), disodium hydrogen phosphate (17549), NaCl (40731), dimethylformamide (17754), and SDS (184190) were procured from Merck Limited. MassRuler DNA ladder mix (S0403) was purchased from Thermo scientific. Nuvia Ni IMAC resin (7800800) was purchased from Biorad. The analytical reagents used were of high purity > 98 %.

2.3 Experimental methods

2.3.1 Competent cell preparation and plasmid DNA transformation

E coli DH5α and BL21 competent cells were transformed with plasmid DNA for DNA isolation and protein expression/purification respectively. The competent cells were prepared using the chemical method by treatment with $CaCl_2$. The starter culture in 5 mL LB (Luria Broth) was initiated with 1 % inoculum and allowed to grow overnight in a shaker incubator at 37 °C, 180 rpm. After overnight growth, 500 µL starter culture volume was used to grow cells in 50 mL LB. The 50 mL culture was grown at 37 °C, 180 rpm until the $Abs_{600\ nm}$ reached 0.4. After incubation at 4 °C for 10 minutes, the culture was centrifuged at 3000 rpm for 10 minutes at 4 °C and the supernatant discarded. The cell pellet was resuspended in 15 mL ice-cold 80 mM $MgCl_2$ and 20 mM $CaCl_2$ by inverting gently. The cells were collected after another round of centrifugation at 3000 rpm for 10 minutes at 4 °C and resuspended in 900 µL $CaCl_2$ + 100 µL autoclaved glycerol. 100 µL aliquots of the prepared competent cells were stored at -80 °C until they were used for transformation. The competent cell preparation was done in sterile conditions, the solutions used were autoclaved before use and the preparation was done in a laminar air flow (LAF).

DH5α competent cells were transformed with plasmid DNA using the heat shock method. 25 ng plasmid DNA was mixed with cells and heat shock given at 42 °C for 60 seconds. The cells were incubated in ice for 2 minutes and then 800 µL autoclaved LB was added. The cells were allowed to grow for 1 hour at 37 °C and 180 rpm in a shaking incubator. The cells were plated in Luria agar plates containing 100 µg/mL ampicillin. The plates were incubated at 37 °C for 12 hours. Single colonies after transformation were picked up and used to start 5 mL LB Amp containing

cultures. Overnight grown cultures were used to perform plasmid isolation using QIAprep spin miniprep kit. Purified plasmid DNA was stored at -20 °C.

2.3.2 Expression and purification of α_3C and α_3W proteins

BL-21 (DE3) cells were transformed with a modified pET32b plasmid containing gene encoding fusion protein for overexpression and purification. After transformation, a single colony was picked from Luria agar plates and used to inoculate 5 mL LB containing 100 µg/mL Amp. The starter culture was used to inoculate large cultures which were grown at 37 °C and 180 rpm. When $Abs_{600\ nm}$ reached 0.6—0.7, protein expression was induced by 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). After protein overexpression for 3 hours, the cells were centrifuged at 6000 rpm for 10 minutes at 4 °C. The pellets were used for purification of fusion protein. The fusion protein consists of thioredoxin tag, His tag, and α_3C/α_3W protein and was purified using the affinity of His tag for Nickel beads according to previously published protocols.[9,75]

The pellet was resuspended in lysis buffer (20 mM Tris, 500 mM NaCl, 1 mM PMSF, pH 8) and kept at 4 °C. The resuspended pellets were sonicated using an ultrasonicator probe and centrifuged at 12000 rpm for 20 minutes at 4 °C. The supernatant containing the fusion protein was collected and filtered before affinity chromatography. 1.5 mL Ni-NTA beads were pre-equilibrated with binding buffer (20 mM Tris, 500 mM NaCl, pH 8) and the supernatant was allowed to bind to the beads in a gravity flow column. The column was placed on an end-to-end rocker at 4 °C for 3 hours to allow complete binding of His-tag containing fusion protein to Ni^{2+} beads. The proteins were eluted using an imidazole gradient (20 mM Tris, 500 mM NaCl, 50—400 mM imidazole, pH 8) and the purer fractions pooled. The concentration determination of pooled α_3C/α_3W fusion protein was done using extinction coefficients of 13980 $M^{-1}cm^{-1}$ and 19580 $M^{-1}cm^{-1}$. The fusion proteins were dialyzed in cleavage buffer (20 mM Tris, 500 mM NaCl, 2.5 mM $CaCl_2$, pH 8) at room temperature and the cleavage was done using thrombin protease. Thrombin was added to the protein in a 1:2000 ratio (w/w) and the cleavage reaction was setup in a dialysis bag for 16 hours at room temperature. After protein cleavage, PMSF was added to the reaction mixture to inhibit protease activity. The cleaved protein of interest was separated from tag using a second round of affinity chromatography. The products of the cleavage reaction were allowed to bind to 1.5 mL Ni-NTA beads pre-equilibrated with cleavage buffer. Subsequent to 3 hours of binding, the flowthrough containing the protein of interest was collected and the purity of cleaved protein was ensured using SDS-PAGE and MALDI-TOF spectroscopy. The proteins HSA, BSA, HEWL, and PRM of high purity purchased from Sigma were used without further purification.

2.3.3 SDS-PAGE electrophoresis

Reducing SDS-PAGE was used to monitor the purity of purified proteins (Figure A3, Appendix). The resolving gel contained 15 % acrylamide as recommended for a ~7.4 kDa protein. The composition of reagents used in the SDS-PAGE electrophoresis are given in the appendix. Sample preparation for electrophoresis was done by mixing samples with gel loading buffer in a 4:1 ratio and heating at 95 °C for 5 minutes. Electrophoresis was carried out in a vertical electrophoresis

system using mini-PROTEAN Tetra setup (Make: BioRad). The run was carried out at 80 V for 2 hours and the gel stained with Coomassie brilliant blue after completion of the run. After sufficient staining, the gels were destained using destaining solution and the images acquired. Standard proteins (in the 14—90 kDa molecular weight range) were run along with samples to give an approximate idea about the molecular mass of the electrophoresed proteins.

2.3.4 MALDI-TOF spectroscopy

The mass of α_3C and α_3W proteins were calculated using the amino acid sequence. This was compared to the mass determined experimentally (Figure A4, Appendix) by a MALDI-TOF mass spectrometer (Make: Bruker Daltonics). A saturated solution of matrix Sinapic acid was prepared in TA-30 solvent. The solution was prepared by vortexing for 30 minutes followed by centrifugation at 12000 rpm for 10 minutes. The supernatant containing matrix was mixed with protein in a 2:1 ratio and spotted on the target plate. The flexControl software was used for data acquisition and analysis was done using flexAnalysis software by Bruker Daltonics.

2.3.5 Absorption spectroscopy

The absorbance spectra were measured in a UV-Vis LAMBDA-25 spectrophotometer (Make: Perkin Elmer). The spectra of proteins used in the study were recorded in the 250—800 nm region at room temperature (~25 °C). The spectra were collected using UV transmissible quartz cuvettes of 1 cm path length. Multiple scans were recorded with 1 nm bandwidth and 480 nm/min scanning speed and the average reported. The experiments were performed for at least three independent samples. The concentration of samples used are mentioned in the figure legends. Absorption spectroscopy was used to measure the extinction coefficient of proteins α_3C, α_3W, and PRM. Absorbance increase with concentration was monitored for α_3C, α_3W, and PRM proteins prepared in deionized water. The buffer systems used in experiments have been specified in the figure legends. The effect of pH and salt on α_3C, α_3W, and PRM absorbance spectra were monitored. The concentration determination of PRM was done using the formula $Abs_{230\ nm} = 1$, when $Conc_{PRM} = 0.47$ mg/mL.[76] The effective nucleotide concentration of Calf Thymus genomic DNA was determined using $\varepsilon_{260\ nm} = 6500\ M^{-1}cm^{-1}$.[77] HSA, HEWL, and BSA concentration determination was done using molar extinction coefficient values of 36500, 38940, and 43284 $M^{-1}cm^{-1}$.[78,79,80] The absorbance spectra of PRM-gDNA, HEWL-gDNA, and BSA-gDNA mixtures were measured. Absorbance was also used to monitor protein unfolding.

2.3.6 α_3C and HSA dansyl labelling

The primary amines in HSA and the single Cys in α_3C were conjugated with extrinsic chromophore Dansyl. IAEDANS (5-({2-[(iodoacetyl)amino]ethyl}amino)naphthalene-1-sulfonic acid) and Dansyl Chloride (5-(DimethylAmino)Naphthalene-1-Sulfonyl chloride) were used for protein labelling reactions. IAEDANS and Dansyl Chloride stock solutions were prepared in dimethylformamide. 250 µM HSA was labelled with 0.06 mM Dansyl Chloride (for primary amine labelling) in 0.1 M bicarbonate buffer (pH 9.3) as described previously.[81] After addition of dye

and stirring at 150 rpm for 90 minutes at 4 °C, protein was eluted as 1 mL fractions in elution buffer (50 mM NaH_2PO_4, pH 7) using PD-10 desalting columns. 200 µM α_3C was labelled with 10-fold excess of IAEDANS (for Cys labelling) in 25 mM Tris, 25 mM NaCl, 5 mM TCEP (pH 7.5).[82] After addition of the dye to protein solution and stirring at 150 rpm for 3 hours at room temperature, protein was eluted in multiple 1 mL fractions with elution buffer (50 mM NaH_2PO_4, pH 7). The labelled protein was separated from the free dye fractions using PD-10 desalting columns. Absorbance of labelled protein and free dye fractions were collected. Correction factor was calculated using the Abs_{280}/Abs_{339} of free dye fractions. Correction factor was used to calculate the corrected absorbance at 280 nm using formula $Abs_{corrected} = Abs_{280} - (C.F.*Abs_{339})$ and ε_{280} was used to calculate protein concentration. Dye concentration was calculated using ε_{339} = 3370 $M^{-1}cm^{-1}$.[81] The efficiency of the protein labelling was reported and protein labelling was also verified using steady-state fluorescence anisotropy r_{ss} of the conjugated dansyl probe.

2.3.7 Steady-state fluorescence and anisotropy measurements

ProCharTS luminescence, dansyl fluorescence, dansyl anisotropy, Trp fluorescence, and Trp anisotropy measurements were made in Fluoromax-4 instrument (Make: Horiba scientific). Trp fluorescence and anisotropy were collected upon excitation at 295 nm and ProCharTS luminescence was monitored upon excitation at 340 nm. The fluorescence and anisotropy of extrinsic chromophore dansyl conjugated to proteins HSA and α_3C was collected with Ex λ = 340 nm. For fluorescence and luminescence experiments the emission collection range, excitation and emission slit widths have been mentioned in the figure legends. The reported spectra were collected using the S1c/R1 mode (corrected for minor changes in excitation light intensity and detector wavelength sensitivity). Contributions from solvent and Raman scattering were subtracted manually. The excitation and emission luminescence spectra of proteins α_3C, α_3W, and PRM were recorded. Luminescence increase with concentration was monitored for α_3C, α_3W, and PRM proteins prepared in deionized water. The buffer systems used in experiments have been specified in the figure legends. The effect of pH and salt on α_3C, α_3W, and PRM luminescence spectra was monitored. The fluorescence spectra of HEWL in HEWL-gDNA mixtures was also measured. ProCharTS luminescence, Trp fluorescence and anisotropy were used to monitor protein unfolding. The measurements were done for at least two independent samples. The experiments were performed at room temperature unless specified (~25 °C).

Stokes shifts at multiple excitation λ were calculated for α_3C, α_3W, and PRM steady-state ProCharTS luminescence spectra using the formula: $\overline{\Delta\nu} = \bar{\nu}_{ex} - \bar{\nu}_{em}$.[15] The Stokes shift was calculated in wavenumbers (cm^{-1}) from excitation wavelength and position of the emission max.

2.3.8 Time-resolved fluorescence and luminescence decays

Time-resolved fluorescence intensity decay measurements were made in a TCSPC instrument using the Delta-Pro equipped with motorized polarizer (Make: Horiba scientific). To measure NATA, α_3C, and α_3W intensity decays, the samples were excited with 295 nm Deltadiode pulsed light source and emission collected using a 340 ± 20 nm emission bandpass filter. The 295 nm

excitation source had a 20 MHz repetition rate and the corresponding IRF FWHM was ~0.8 ns. α_3C luminescence intensity decays were additionally collected upon excitation at 295 and 340 nm using 320 and 370 nm emission longpass filters.

For all the measurements made, the emission was collected up to 15000 counts. The emission polarizer was kept at 54.7 ° to avoid contributions from rotational diffusion. The decays were collected in 2202 channels with a time resolution of 28 ps/channel. It was ensured that inner filter effects were avoided during data acquisition by keeping absorbance of sample at excitation λ less than 0.1. Intensity decays were collected for three independently prepared samples and data analysed. The data analysis was done using software developed by Dr. N. Periasamy, Tata Institute of Fundamental Research (TIFR), Mumbai. The discrete analysis was performed by the method of nonlinear least squares (NLLS) [36,72] and the fluorescence lifetime distribution analysis was done by maximum entropy method (MEM).[36,73] The goodness of fit was analysed by χ^2 and residual distribution.

2.3.9 Circular dichroism spectroscopy

The CD spectra measurements of DNA and proteins in multiple conditions were performed in a Jasco J-1500 instrument using a 2 mm pathlength quartz cuvette (Make: Starna Scientific). CD measurements were made in the 190—260 nm region or the 190—300 nm region. Additional parameters included the use of 0.1 nm data pitch, 2 nm bandwidth, and 100 nm/min scan speed. Blank subtraction from CD spectra of the samples was done manually, data was averaged and reported. The concentration of samples and the buffer conditions are mentioned in the legends. The experiments were repeated for two independent samples. The measurements were made at room temperature unless specified (~25 °C). The ellipticity (mdeg) recorded in the CD measurements were converted to mean residual ellipticity (deg cm^2 $dmol^{-1}$) when protein unfolding by chemical denaturant was analysed using $MRE_{222\ nm}$:

$$MRE_{222} = \frac{\theta_{222} * MRW}{10 * l * c} \qquad 2.29$$

here l = path length of cuvette in cm, c = concentration in mg/ml, and MRW = molecular weight/number of peptide bonds. Secondary structure analysis was done using K2D3 program.[83]

2.3.10 HSA, α_3C, and α_3W protein unfolding

Protein unfolding was monitored using conventional methods and ProCharTS. HSA was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at room temperature (~25 °C). ProCharTS absorbance, MRE_{222}, ProCharTS luminescence, Trp fluorescence, Trp anisotropy, Dansyl fluorescence and Dansyl anisotropy values were used to characterize HSA unfolding by the chemical denaturant GdnHCl. The concentration of denaturant stock solutions were determined using the refractive index measured in an Abbe's refractometer.[84] α_3W was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, pH 7.2) for 12 hours at room temperature (~25 °C). ProCharTS absorbance, MRE_{222}, Trp fluorescence and anisotropy values at multiple points in

the protein unfolding pathway were used to characterize α_3W unfolding. α_3C was incubated in 0—6 M GdnHCl (20 mM Tris, pH 7.2) for 12 hours at room temperature (~25 °C). ProCharTS absorbance, MRE_{222}, Dansyl fluorescence and anisotropy values at multiple points in α_3C unfolding pathway were used to characterize unfolding.[85]

Quantitative analysis of protein unfolding was done by estimating the denaturation midpoint of proteins HSA, α_3C, and α_3W. Denaturation midpoint is the denaturant concentration at which folded and unfolded state of proteins are equally populated. The denaturation midpoint was obtained by using the nonlinear regression analysis module in Origin software. The choice of model depended on the R^2 values obtained and the standard error in predicting denaturation midpoint. The denaturation midpoint, standard error, and R^2 of the best fitted model was reported.

2.3.11 DNA condensation assay

The protein, DNA sample preparation, and the condensation assay were performed in 50 mM NaCl. To a fixed concentration of protein (PRM/HEWL/BSA), an increasing amount of calf thymus gDNA was added. Protein and DNA mixing was done at 4 °C for 15 minutes at 120 rpm. After DNA-protein mixing, the samples were centrifuged at 12000 rpm for 15 minutes and the supernatant was used for absorbance, CD, and agarose gel electrophoresis experiments. PRM ProCharTS, BSA absorbance, HEWL absorbance and fluorescence were used to characterize DNA-protein binding.

In a separate methodology to monitor DNA-protein binding, increasing concentration of protein was added to a fixed concentration of gDNA (calculated in effective nucleotide concentration). Protein and DNA mixing was done at 4 °C for 15 minutes at 120 rpm. After DNA-protein mixing, the samples were centrifuged at 12000 rpm for 15 minutes and the supernatant collected.[7,86] gDNA absorbance, CD, and agarose gel electrophoresis of the supernatant was used to characterize DNA-protein binding. The scattering at 350 nm of DNA-protein mixtures (prior to centrifugation) was also used to characterize DNA-protein binding.

2.3.12 Electrophoretic mobility shift assays

The protein and DNA sample preparation was done in 50 mM NaCl. To a fixed concentration of gDNA (calculated in effective nucleotide concentration), increasing concentrations of protein were added. The gDNA-protein mixture was allowed to bind for 15 minutes. After completion of binding, the change in electrophoretic mobility of DNA-protein complex and condensation of DNA by protein was studied using agarose gel electrophoresis.[18] The precipitation of nucleoprotein complex upon gDNA binding with HEWL and PRM was also confirmed by the absence of DNA in DNA-protein mixture supernatants using agarose gel electrophoresis. Ethidium bromide was used to visualize DNA after gel electrophoresis.

CHAPTER 3

OBSERVING PROTEIN CHARGE TRANSFER SPECTRA IN PROTEINS α_3C, α_3W, AND PRM

3.1 ProCharTS in α_3C, α_3W, and PRM

Electronic absorption and luminescence in the near UV-Visible region was observed in proteins α_3C, α_3W, and PRM. The novel absorbance observed was attributed to the phenomenon of protein charge transfer involving the charged amino acid sidechains and the peptide backbone. The luminescence observed upon excitation with near UV wavelengths originated due to charge recombination luminescence. The luminescence quantum yield was low, possibly due to poor radiative recombination. The linear increase in ProCharTS absorbance and luminescence with concentration was established. The above results are described in the following sections in detail.

3.2 Primary structure of α_3C, α_3W, and PRM

The proteins α_3C, α_3W, and PRM are rich in charged amino acids (Figure 3.1). α_3C and α_3W are natively folded three helical bundle proteins whereas salmon sperm protamine is intrinsically disordered. In α_3C and α_3W, numerous charged amino acid residues in close contact (within 10 Å) have been observed in the PDB structures.[87,88] TDDFT calculations performed on α_3C were used to validate the presence of protein charge transfer in α_3C by Prasad *et al.*[9] α_3W is the single Trp variant of α_3C, where Cys-34 has been replaced by Trp (the other amino acid residues are identical). α_3C and α_3W are rich in Lys and Glu whereas PRM is rich in Arg (Figure 3.1).

> PDB ID 2LXY α_3C

GSRVKALEEK VKALEEKVKA LGGGGRIEEL KKKCEELKKK
IEELGGGGEV KKVEEEVKKL EEEIKKL

> PDB ID 1LQ7 α_3W

GSRVKALEEK VKALEEKVKA LGGGGRIEEL KKKWEELKKK
IEELGGGGEV KKVEEEVKKL EEEIKKL

> P69014 Salmon sperm protamine

PRRRRSSSRP VRRRRRPRVS RRRRRRGGRR RR

Figure 3.1: The primary structure of α_3C, α_3W, and PRM. The UniProt ID of PRM and PDB IDs of α_3C and α_3W are used to identify the three proteins.[87,88] The amino acid sequences with the charged amino acids highlighted in red and Cys/Trp highlighted in light blue.

α_3C and α_3W contain 17 Lys, 17 Glu, and 2 Arg out of 67 amino acid residues. PRM contains 21 Arg out of 32 amino acid residues in sequence. Other charged amino acid residues are absent in PRM. α_3W contains 1 Trp whereas α_3C and PRM are devoid of aromatic amino acids. Abundance of charged amino acids in α_3C, α_3W, and PRM (Table 3.1) implied the possibility of extensive charged amino acid sidechain contact formation. The occurrence of these contacts are expected to be higher especially in the folded proteins α_3C and α_3W.

Protein	No of Positively charged amino acids (% +ve charge)	No of Negatively charged amino acids (% -ve charge)	Total charge content (%)
α_3C	19 (28.4)	17 (25.3)	53.7
α_3W	19 (28.4)	17 (25.3)	53.7
PRM	21 (63.6)	0	63.6

Table 3.1: Charged amino acid content of α_3C, α_3W, and PRM. The richness of charged amino

acid residues in α_3C, α_3W, and PRM are reflected in the percent of charged amino acid residues in sequence.

3.3 ProCharTS absorbance in α_3C, α_3W, and PRM

α_3C absorbed strongly in the near UV region ($\varepsilon_{250\ nm}$ = 7171.3 $M^{-1}cm^{-1}$) with a long tail extending into the visible region ($\varepsilon_{800\ nm}$ = 400.3 $M^{-1}cm^{-1}$) shown in Figure 3.2A. α_3W absorbance comprised of Trp absorbance (250—325 nm) and CT absorbance in the near UV and visible region of the electromagnetic spectrum ($\varepsilon_{800\ nm}$ = 383.5 $M^{-1}cm^{-1}$) shown in Figure 3.2B. Previous TDDFT computational studies discovered extensive peptide backbone to sidechain and sidechain to sidechain charge transfer transitions in protein α_3C[9] which explains the observed novel α_3C and α_3W absorbance.

PRM showed moderate absorbance in the near UV region even though it lacked aromatic amino acids (Figure 3.2C). The observed absorbance was attributed to PBS-CT involving the Arg sidechain (64 % Arg content) and peptide backbone. PRM absorbance was observed from 250 nm (ε = 752.1 $M^{-1}cm^{-1}$) till 350 nm (ε = 13.2 $M^{-1}cm^{-1}$). The long tail observed in α_3C and α_3W absorbance was absent from PRM absorbance, possibly because of the absence of SS-CT transitions (PRM lacks oppositely charged amino acid residues). Another reason for low PRM extinction coefficient is the intrinsically disordered nature of PRM. In contrast to a folded protein, long distance Arg sidechain contact formation would be lower due to the absence of the protein fold which brings together distant parts of the polypeptide chain in close contact. Additionally, α_3C, α_3W, and PRM absorbance showed poor overlap with simulated Rayleigh scatter ruling out that the observed absorbance was due to scattering.

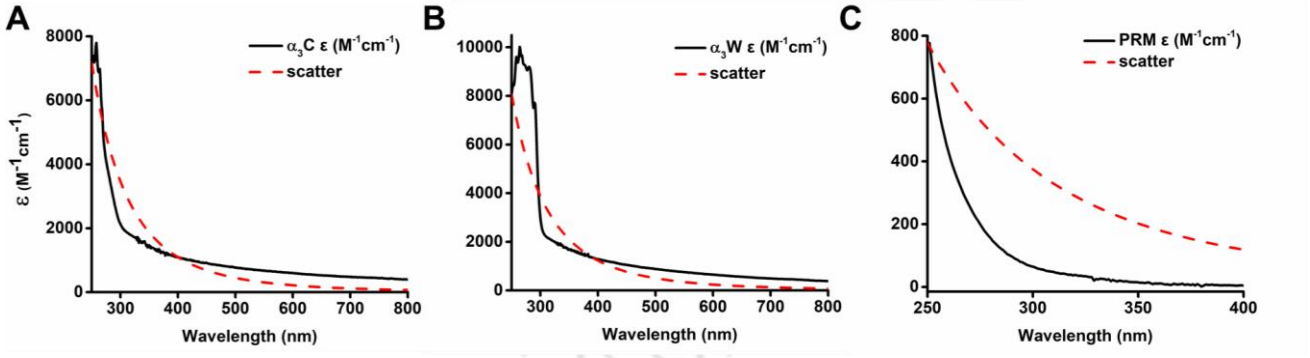


Figure 3.2: Molar extinction coefficient of proteins α_3C, α_3W, and PRM. The concentration of α_3C and α_3W were determined by Lowry's method[89] and the concentration of PRM was determined by far UV method.[90] The extinction coefficient values were calculated from measured absorbance for α_3C (A) and α_3W (B) between 250—800 nm and between 250—400 nm for PRM (C). The simulated Rayleigh scatter was calculated from absorbance at 250 nm (based on $1/\lambda^4$ dependence) and plotted along with the extinction coefficient of the proteins. The samples were prepared in deionized water.

The extinction coefficient values of α_3C, α_3W, and PRM at selected wavelengths are reported in Table 3.2. α_3C and α_3W ProCharTS overlapped in the 325—800 nm region. PRM had negligible absorbance beyond 350 nm.

	250	300	350	400	500	600	700
α_3C ε ($M^{-1}cm^{-1}$)	7171	2138	1474	1088	775	600	481
STDEV	122	11	8	11	8	19	32
α_3W ε ($M^{-1}cm^{-1}$)	8052	2991	1684	1281	882	648	493
STDEV	30	58	61	22	1	24	31

	250	260	270	280	295	320	350
PRM ε ($M^{-1}cm^{-1}$)	752	397	239	141	69	37	13
STDEV	6	3	2	2	11	7	10

Table 3.2: Molar extinction coefficient values of proteins α_3C, α_3W, and PRM. The concentration of α_3C and α_3W were determined by Lowry's method[89] and the concentration of PRM was determined by far UV method.[90] The extinction coefficient values and corresponding standard deviation were tabulated.

3.4 ProCharTS luminescence in α_3C, α_3W, and PRM

α_3C and α_3W samples showed luminescence upon excitation with multiple wavelengths in the 295—410 nm region (Figure 3.3A and 3.3D). α_3W luminescence was measured upon excitation in the 325—410 nm region to avoid excitation of the Trp chromophore. The observed α_3C and α_3W novel luminescence were characterized by moderate Stokes shifts (3090 to 6136 cm^{-1}) shown in Figure 3.3C and 3.3F. α_3C and α_3W ProCharTS luminescence did not obey Kasha's rule and excitation wavelength dependent shift in emission maxima was observed. The shift in emission maxima with excitation wavelength hinted at the possible involvement of multiple CT states in the origin of α_3C and α_3W luminescence. Area under the emission curves decreased when longer excitation wavelengths were used to illuminate the samples (Figure 3.3B and 3.3E). Negligible luminescence was obtained from both α_3C and α_3W upon excitation with wavelengths longer than 410 nm.

Significantly red-shifted emission was observed from PRM upon excitation with wavelengths in the 250—320 nm region with an emission maximum at ~424 nm (Figure 3.3G). The luminescence obtained from PRM was characterized by high Stokes shifts (7609—16415 cm^{-1}) shown in Figure 3.3I. PRM emission adhered to the Kasha's rule and emission maxima were independent of excitation λ used. Maximal emission was obtained upon excitation at 260 nm which was also reflected in the area under luminescence curve (Figure 3.3H).

The observed novel intrinsic α_3C, α_3W, and PRM luminescence was attributed to charge recombination luminescence as illustrated by Kumar *et al*.[15]

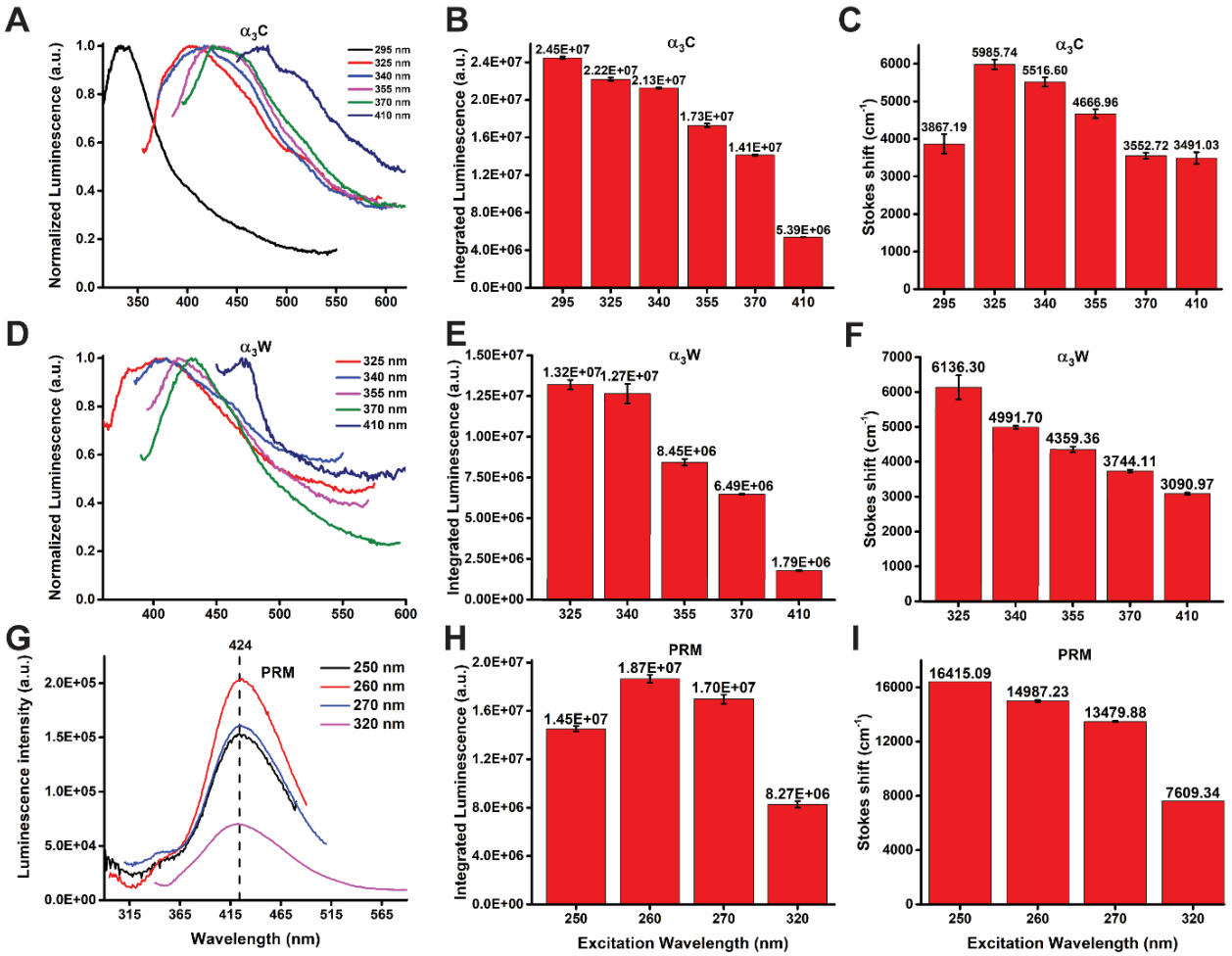


Figure 3.3: Luminescence spectra of proteins used in the study. 40 µM α_3C (A), 20 µM α_3W (D), and 75 µM PRM (G) emission spectra upon excitation at multiple wavelengths. The luminescence spectra of α_3C and α_3W were normalized to 1. The area under the luminescence curve for α_3C (B), α_3W (E), and PRM (H) at multiple excitation wavelengths and the corresponding Stokes shifts for α_3C (C), α_3W (F), and PRM (I). The spectra were collected using Ex. slit = 5 nm and Em. slit = 15 nm. The samples were prepared in deionized water.

3.5 ProCharTS luminescence quantum yields

The low to moderate quantum yields of multiple monomeric proteins rich in charged amino acids were reported at 355 nm.[15] The luminescence quantum yields of α_3C and α_3W measured at 355 nm were 0.0040 and 0.0037 (Table 3.3). The observed quantum yields were low and similar to each other. The quantum yields were calculated with reference to DPA whose quantum yield is known at 355 nm. DPA was chosen as reference due to overlap between emission profile of α_3C, α_3W luminescence, and DPA fluorescence.

The absorbance and emission profile of PRM is unique with the observation of significantly red-shifted emission upon excitation with near UV wavelengths. Indole quantum yield is known at 270 nm and was used to calculate PRM luminescence quantum yield. There is a slight overlap between

the absorption spectra of PRM and indole but the emission spectra are different. PRM was found to have a moderate quantum yield of 0.0108 at 270 nm (Table 3.3).

The low quantum yields observed for ProCharTS luminescence implied that the majority of absorbed photons were de-excited by non-radiative pathways. The phenomenon of ProCharTS luminescence is less well understood and possible reason for low quantum yield of α_3C, α_3W, and PRM is likely to be the poor radiative recombination between CT states.

SAMPLE	Int Fl. Intensity	Absorbance	Absorptance	Ref Index	$QY_{355\ nm}$	$QY_{270\ nm}$
DPA	4.74E+08	0.025	0.055	1.426		
α_3C	1.17E+06	0.026	0.058	1.333	0.0040	
α_3W	1.03E+06	0.025	0.056	1.333	0.0037	
Indole	1.91E+09	0.025	0.055	1.332		
PRM	2.31E+07	0.022	0.050	1.334		0.0108

Table 3.3: α_3C, α_3W, and PRM luminescence quantum yields. α_3C and α_3W quantum yields were calculated at 355 nm using DPA as a reference. PRM quantum yield was calculated at 270 nm using indole as a reference. The quantum yields were calculated using the following formula: $\phi_f^i = \phi_f^s(F^i f_s \eta_i^2 / F^s f_i \eta_s^2)$ where ϕ_f^s and ϕ_f^i are quantum yields, F^s and F^i are integrated fluorescence intensities, f_s and f_i are absorptance, η_s^2 and η_i^2 are refractive index squared of standard and sample respectively. The absorptance values f_x were calculated using $f_x = 1 - 10^{-Ax}$.[91] For DPA, α_3C, and α_3W, the Ex. slit was 1 nm and Em. slit was 5 nm. For PRM and indole, the excitation and emission slit widths were 2 and 15 nm respectively. DPA was dissolved in cyclohexane and the other samples were prepared in deionized water.

3.6 ProCharTS luminescence excitation spectra

The excitation spectra of α_3C, α_3W, and PRM for the observed luminescence were collected. The excitation spectra consisted of two peaks and the excitation maxima was between 260—270 nm for α_3C and PRM (Figure 3.4A and 3.4C). Due to the high quantum yield of Trp, this peak is masked in the α_3W excitation spectra and only the Trp excitation appears prominently in the 250—315 nm region (Figure 3.4B). In α_3W, contribution from both Trp fluorescence and ProCharTS luminescence was obtained. A second peak of lower intensity corresponding to ProCharTS luminescence with excitation maximum around 360 nm was observed. This peak appeared prominently in the α_3C and PRM excitation spectra. The second peak appeared at ~360 nm in the α_3C excitation spectra and at 320 nm in the PRM spectra. The observation of two peaks in the excitation spectra illustrated that the observed ProCharTS luminescence did not arise from a single state and multiple CT states may be contributing to the observed luminescence.

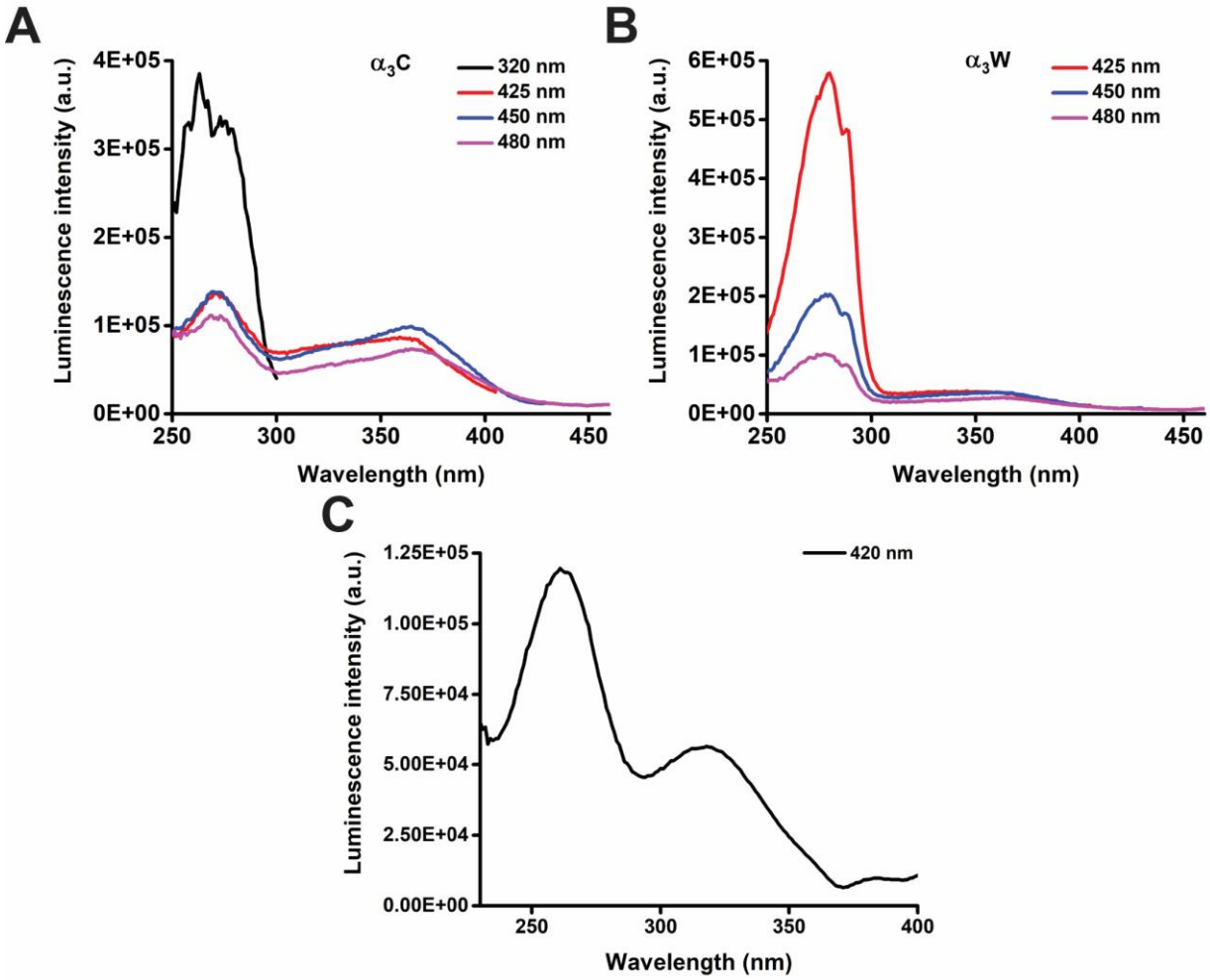


Figure 3.4: Luminescence excitation spectra in α_3C, α_3W, and PRM. 12 µM α_3C (A), 12 µM α_3W (B), and 75 µM PRM (C) excitation spectra was collected at multiple wavelengths. The spectra were collected using Ex. slit = 2 nm and Em. slit = 15 nm. The samples were prepared in deionized water.

3.7 The effect of increasing concentration on α_3C, α_3W, and PRM absorbance

The increase in α_3C, α_3W, and PRM absorbance with concentration was monitored. Significant absorbance from α_3C and α_3W throughout the near UV-Visible region was observed in the 10—50 µM concentration range (Figure 3.5A and 3.5B). 25—150 µM PRM absorbed moderately in the 250—350 nm region (Figure 3.5C). The increase in protein absorbance with concentration was observed visually and the linearity investigated by fitting to a linear function (results plotted in Figure 3.6).

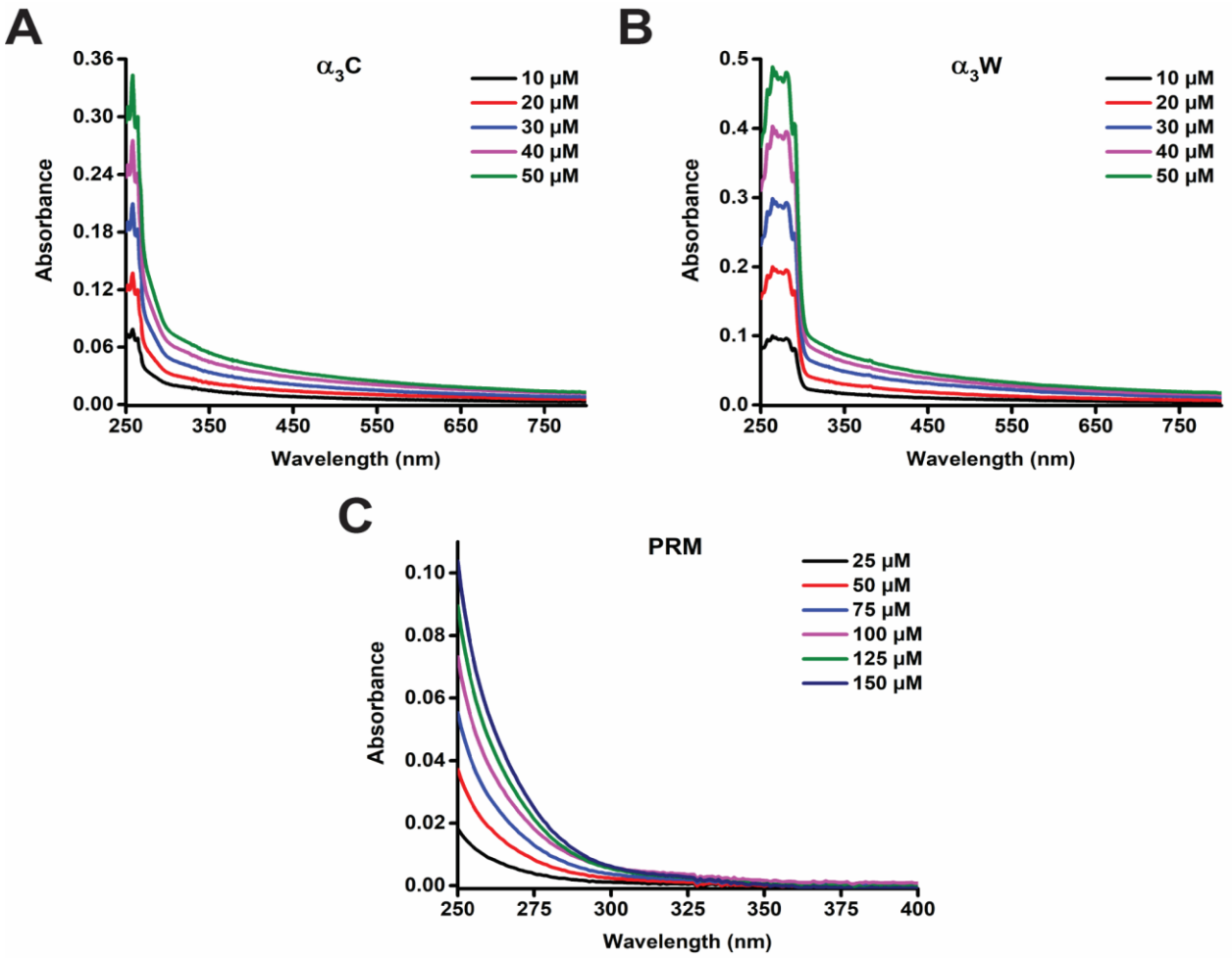


Figure 3.5: Monitoring the increase in α3C, α3W, and PRM ProCharTS absorbance with concentration. α3C (A) and α3W (B) absorbance was collected in the 250—800 nm region for 10—50 μM protein concentration. PRM absorbance (C) was measured in the 250—400 nm region for 25—150 μM protein. The samples were prepared in deionized water.

α3C, α3W, and PRM absorbance increased linearly with concentration in accordance with the Beer-Lambert's law (Figure 3.6). The linearity of absorbance increase was monitored at multiple wavelengths and R^2 values greater than 0.994 were obtained for all fits of absorbance to linear function. The wavelengths monitored, slope, and R^2 values are reported in Table 3.4. The linearity suggested that the observed novel absorbance originated from the monomeric protein and ruled out protein aggregation at higher concentrations. Protein aggregation would have led to increased intermolecular charged amino acid sidechain contact formation resulting in greater protein charge transfer. The linear increase in absorbance with concentration thus implied the absence of protein aggregation with increasing protein concentration.

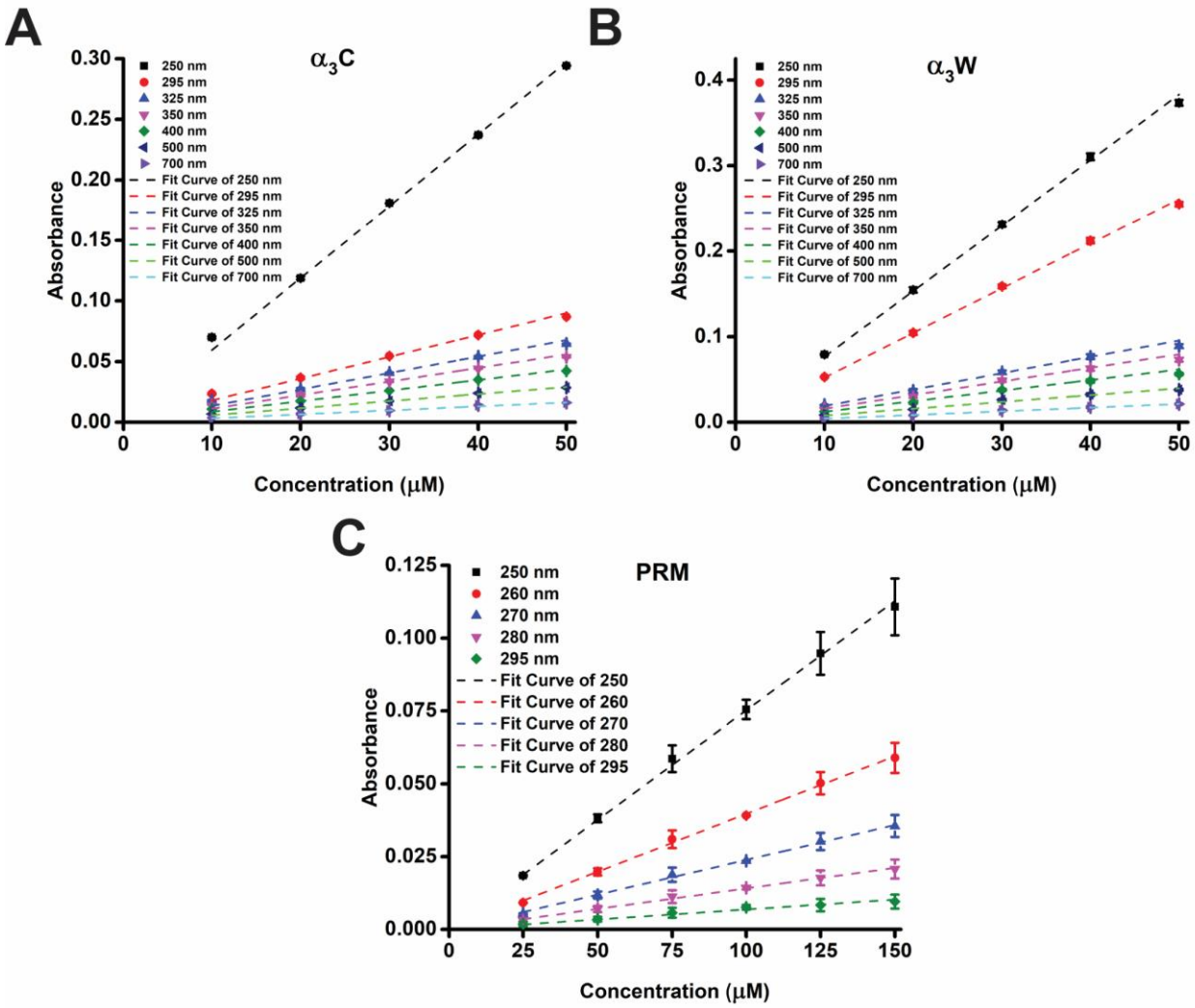


Figure 3.6: Monitoring the linearity in α3C, α3W, and PRM absorbance increase with concentration. α3C (A) and α3W (B) absorbance was collected in the 250—800 nm region for 10—50 μM protein concentration. PRM absorbance (C) was measured in the 250—400 nm region for 25—150 μM protein. The absorbance increase with concentration was fit to a linear function at multiple wavelengths.

Linearity of α3C and α3W ProCharTS absorbance increase was monitored at multiple points in the 250—700 nm region. PRM absorbance linearity was monitored in the 250—295 nm region and reported in Table 3.4. ProCharTS absorbance increase with concentration did not deviate from the linear function at higher protein concentrations. The slope values obtained for α3W absorbance increase were in good agreement with molar extinction coefficient values measured by Lowry's method (Table 3.2). The slope values for α3C were lower than measured extinction coefficients but followed a similar trend of significant absorbance in the near UV-Visible region with a long tail extending well into the visible region.

Sample	Wavelength (nm)	250	295	325	350	400	500	700
α_3C	Slope ($M^{-1}cm^{-1}$)	5940	1800	1350	1120	872	582	327
	R^2	0.9997	0.9992	0.9995	0.9940	0.9993	0.9985	0.9980
α_3W	Slope ($M^{-1}cm^{-1}$)	7660	5220	1910	1590	1230	792	428
	R^2	0.9994	0.9996	0.9967	0.9971	0.9975	0.9956	0.9981
	Wavelength (nm)	250	260	270	280	295		
PRM	Slope ($M^{-1}cm^{-1}$)	752	397	239	141	69		
	R^2	0.9997	0.9997	0.9995	0.9993	0.9945		

Table 3.4: Parameters obtained upon fitting absorbance increase with concentration to linear function. The slope at multiple wavelengths and R^2 values obtained upon fitting absorbance values to linear equation was reported.

3.8 The effect of increasing concentration on α3C, α3W, and PRM luminescence

α_3C and α_3W luminescence was measured upon excitation at 355 nm (Figure 3.7A and 3.7B), and PRM luminescence was measured at 260 nm with increasing protein concentration (Figure 3.7C).

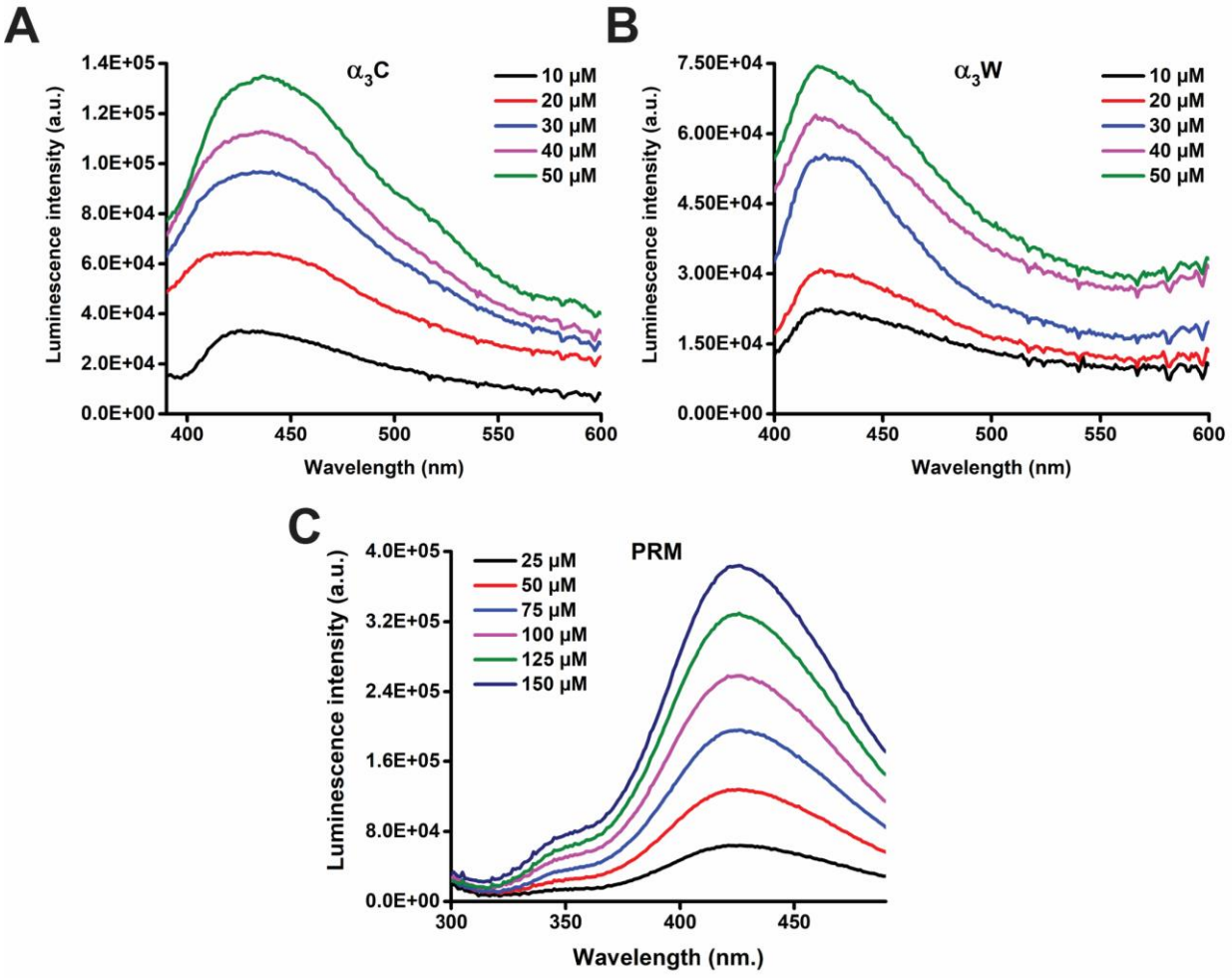


Figure 3.7: Monitoring the increase in α3C, α3W, and PRM ProCharTS luminescence with concentration. α_3C (A) and α_3W (B) luminescence upon excitation at 355 nm was collected in the 390—600 nm region. PRM luminescence (C) was measured in the 300—490 nm region upon

excitation at 260 nm. α_3C and α_3W luminescence was collected for 10—50 µM protein and PRM luminescence was measured for 25—150 µM protein. The samples were prepared in deionized water.

The excitation wavelength was chosen as significant luminescence from α_3C and α_3W at these wavelengths had been observed. Also, the excitation λ in the case of α_3W was distant from the regions in which Trp electronic absorption is expected. Thus the observed luminescence was attributable only to protein charge transfer luminescence. Increase in ProCharTS luminescence was observed with protein concentration. The integrated luminescence intensity (area under the curve) with increasing concentration was fitted to a linear function and a linear increase with concentration was observed (Figure 3.8).

α_3C, α_3W, and PRM luminescence increased linearly with concentration (Figure 3.8). The linearity was confirmed by the R^2 values (greater than 0.994) obtained upon fitting integrated luminescence intensity with increasing concentration to a linear function. The slope and R^2 values of the fits are reported in Table 3.5. The linear increase in luminescence with protein concentration ruled out the possibility of protein aggregates at higher concentration contributing to ProCharTS luminescence.

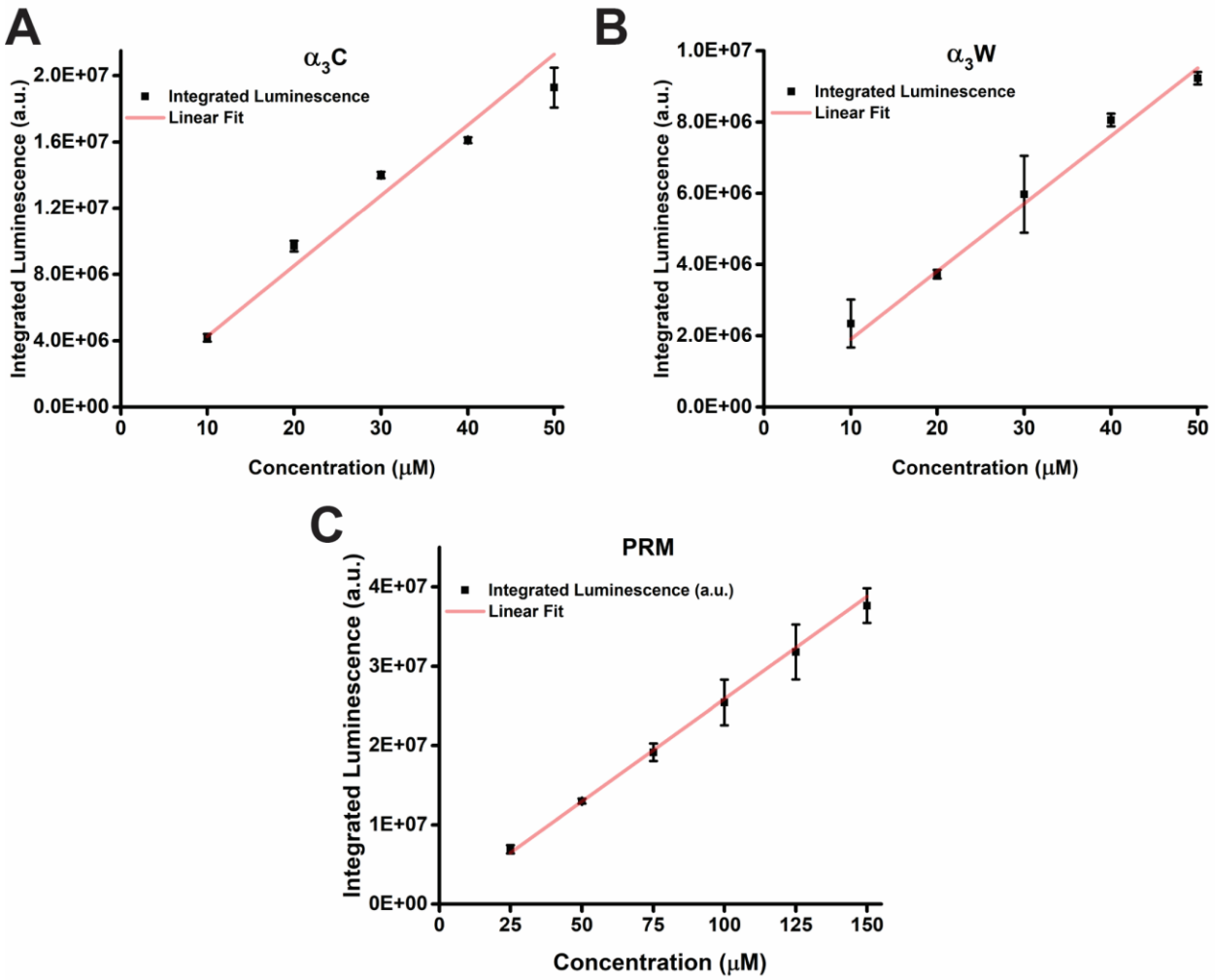


Figure 3.8: Monitoring the linearity in α_3C, α_3W, and PRM ProCharTS luminescence increase with concentration. α_3C (A) and α_3W (B) luminescence upon excitation at 355 nm and

PRM luminescence (C) at 260 nm was monitored with increasing protein concentration. The luminescence intensity increase with concentration was fit to a linear function.

The slope of α_3C, α_3W, and PRM luminescence increase with concentration was tabulated in Table 3.5. The linearity of ProCharTS luminescence increase was reflected in the high R^2 values obtained upon fitting luminescence values to linear function.

Sample	Slope	R^2
α_3C	4.26E+05	0.9943
α_3W	1.90E+05	0.9983
PRM	2.59E+05	0.9996

Table 3.5: Parameters obtained upon fitting luminescence increase with concentration to linear function. The slope and R^2 values obtained upon fitting luminescence values to linear equation was reported.

3.9 α_3C ProCharTS luminescence decay spectra

α_3C time-resolved luminescence intensity decays were collected upon excitation at 295 and 340 nm using a TCSPC instrument. The IRF, measured decay, and fitted data were plotted in Figure 3.9A and 3.9B. The obtained luminescence upon excitation at 295 nm was more intense than at 340 nm (in agreement with steady-state luminescence intensities). NLLS analysis was performed to obtain fit parameters. The decays fit to a sum of three exponentials. The randomness in residual distribution and acceptable χ_R^2 signified that the obtained decays fit the assumed three exponential model. The data fitting was better for intensity decay collected upon excitation at 295 nm than 340 nm (better residual distribution in Figure 3.9C compared to Figure 3.9D). Due to lower luminescence at 340 nm compared to 295 nm, the decay acquisition time was larger when λ_{ex} was 340 nm. Longer acquisition times increased the background counts resulting in higher χ^2, especially in the MEM fit at 340 nm (Figure 3.10D). Due to low PRM ProCharTS extinction coefficient at 295 nm, PRM luminescence intensity decays were not measured.

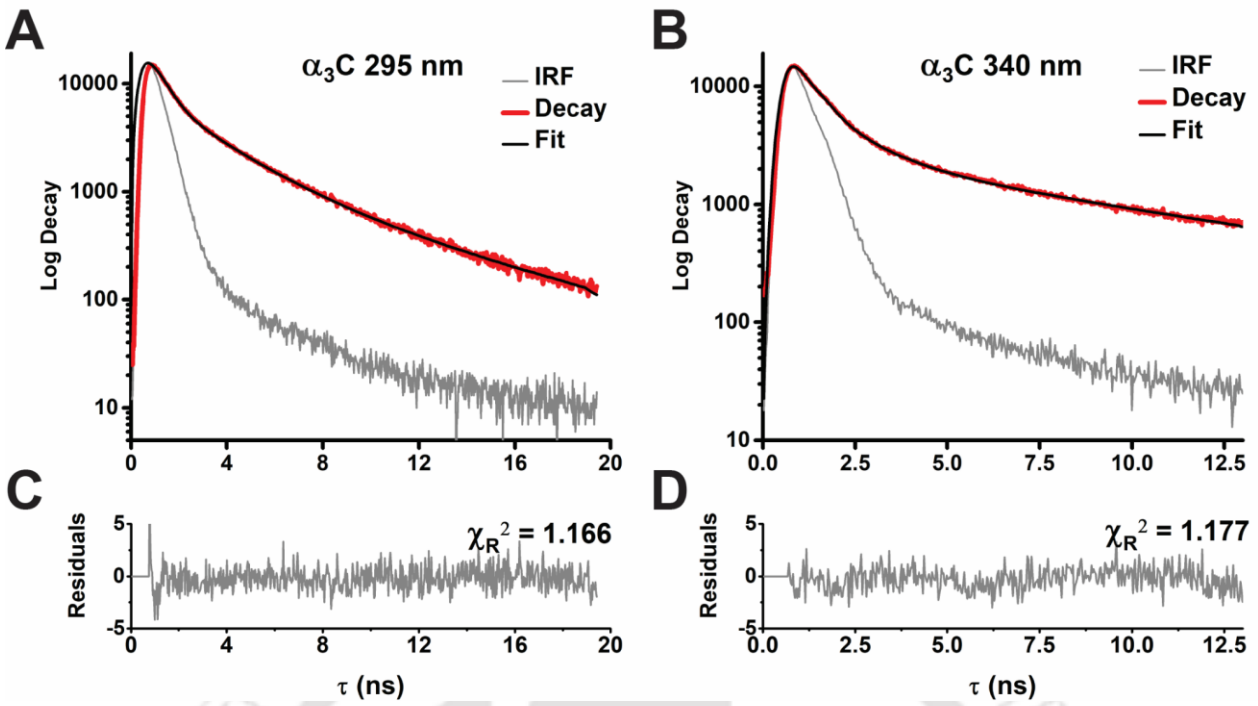


Figure 3.9: α_3C ProCharTS luminescence decays collected upon excitation at 295 and 340 nm. 30 µM α_3C was prepared in deionized water and luminescence decays were collected upon excitation at 295 and 340 nm using 320 and 370 nm longpass filters. The IRF, measured decay, and fitted data for excitation at 295 and 340 nm were plotted in Figure 3.9A and 3.9B. α_3C luminescence decays fit to a sum of three exponentials. The residuals and corresponding χ_R^2 values were reported (Figure 3.9C and 3.9D).

NLLS analysis of collected α_3C luminescence decays was used to obtain fit parameters reported in Table 3.6. The individual components in the three exponential fits and their amplitudes were obtained. The mean lifetime at 295 nm excitation was 0.853 ns and the lifetime at 340 nm excitation was 0.380 ns. The short component had the maximal contribution with α_1 values of 0.822 and 0.914 respectively and the long component had small amplitude values. τ_2 and α_2 were higher upon excitation at 295 nm compared to 340 nm, contributing to the higher mean lifetime observed. The fit quality was acceptable with χ_R^2 lower than 1.2.

Sample	τ_1 (ns)	α_1	τ_2 (ns)	α_2	τ_3 (ns)	α_3	τ_M (ns)	χ_R^2
α_3C 295 nm	0.394 {0.055}	0.822	2.335 {0.055}	0.143	6.150 {0.581}	0.035	0.853 {0.078}	1.166
α_3C 340 nm	0.117 {0.015}	0.914	1.259 {0.070}	0.063	8.636 {0.550}	0.023	0.380 {0.048}	1.177

Table 3.6: α_3C ProCharTS luminescence decay parameters. Parameters obtained upon NLLS analysis of α_3C luminescence intensity decays collected by excitation at 295 and 340 nm. The individual lifetimes, amplitudes, mean lifetime, and χ_R^2 obtained after fitting α_3C luminescence decay to a sum of three exponentials model was tabulated. The standard deviation values were reported in parenthesis. The reported τ values are average of three independent measurements, while α values were taken from the best fitted data.

The obtained decays were analysed by maximum entropy method to obtain the lifetime distribution information. The data obtained from MEM analysis revealed the presence of three lifetime components in agreement with discrete fits. The short component had a larger amplitude and the long component had the smallest amplitude. The amplitude values were normalized and plotted in Figure 3.10. The residuals for fit at 295 nm were random and the χ^2 was acceptable. Randomness in residual distribution was also obtained at 340 nm, but χ^2 of 1.301 was high. The lifetime distribution analysis pointed to the possibility of multiple CT states being involved in the observed phenomenon rather than the luminescence coming from a single state. The measurement and analysis of time-resolved luminescence decays were thus used to characterize α_3C ProCharTS luminescence.

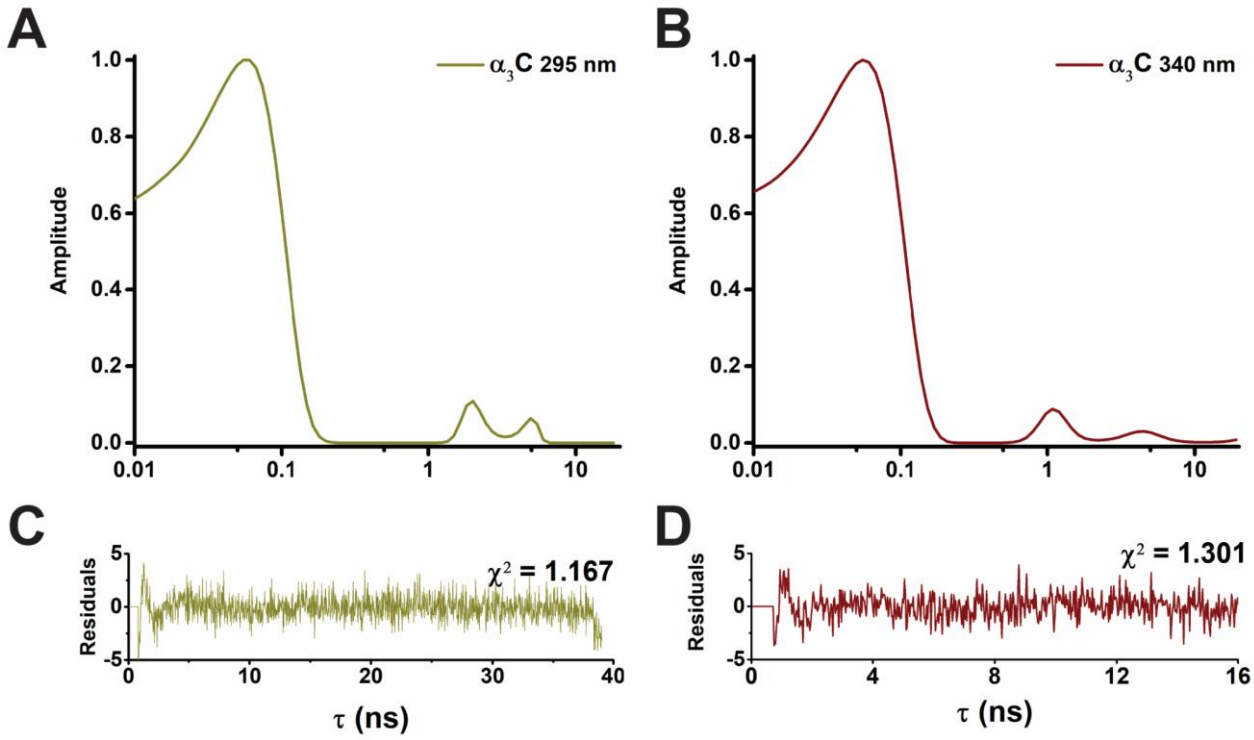


Figure 3.10: α_3C ProCharTS luminescence lifetime distribution. 30 µM α_3C was prepared in deionized water and luminescence decays were collected upon excitation at 295 and 340 nm using 320 and 370 nm longpass filters and analysed by MEM. The normalized amplitudes for lifetime distribution were plotted with log (time) in Figure 3.10A and 3.10B. The residuals and corresponding χ^2 values were also reported (Figure 3.10C and 3.10D).

3.10 Conclusions

1. The presence of ProCharTS in α_3C, α_3W, and PRM was attributed to the abundance of charged amino acids constituting the protein.
2. The strong absorbance by α_3C and α_3W in the near UV-Visible region was possibly due to a combination of extensive PBS-CT and SS-CT transitions. Only PBS-CT was possible in the case of PRM, leading to moderate absorbance in the near UV region.
3. α_3C, α_3W, and PRM emission spectra were characterized by moderate to high Stokes shifts and low quantum yields.
4. The absence of excitation wavelength dependent shift in PRM emission maxima and differences compared to α_3C/α_3W emission were possibly due to the absence of SS-CT transitions. The differences in the observed ProCharTS luminescence highlighted that the phenomenon was less well understood than ProCharTS absorbance.
5. The linear increase in ProCharTS absorbance and luminescence with concentration ruled out aggregation induced effects at higher protein concentrations.
6. Multi-exponential decay kinetics of α_3C luminescence pointed to the involvement of multiple CT states in the observed luminescence.

CHAPTER 4

MONITORING THE EFFECT OF SOLVENT PARAMETERS ON α_3C, α_3W, PRM PROCHARTS AND INFLUENCE OF PROCHARTS LUMINESCENCE ON TRP FLUORESCENCE

4.1 Monitoring the effect of solvent parameters on α_3C, α_3W, and PRM ProCharTS

Charged amino acid sidechains in proteins are involved in Protein Charge Transfer Spectra. The protein fold is sensitive to solvent parameters and extreme changes in pH will affect the number of sidechains in close contact. To evaluate how sensitive ProCharTS is to this change in solvent parameters (pH and salt concentration), α_3C, α_3W, and PRM ProCharTS were measured with pH change and in the presence of salt.

4.1.1 The effect of pH extremes on α_3C, α_3W, and PRM ProCharTS

To understand the role of charged amino acid sidechain ionization status on α_3C, α_3W, and PRM ProCharTS, the absorbance, luminescence, and CD spectra were measured at extreme pH.

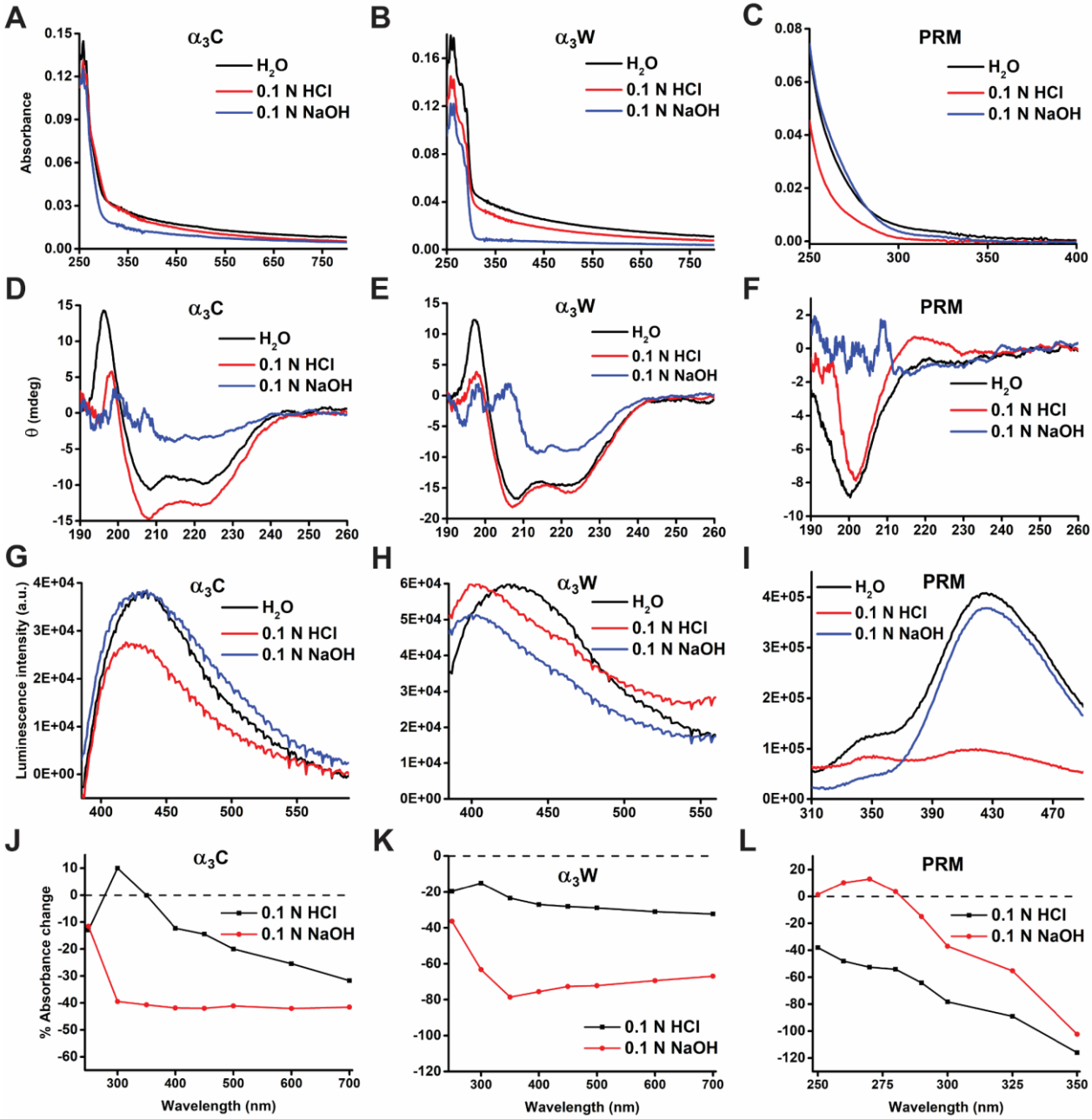


Figure 4.1: Monitoring the effect of extreme pH on α_3C, α_3W, PRM ProCharTS and secondary structure. The absorbance spectra of 20 µM α_3C (A), 20 µM α_3W (B), and 100 µM

PRM (C) were measured in deionized water, at pH 1 (0.1 N HCl), and pH 13 (0.1 N NaOH). The CD spectra between 190—260 nm was recorded in the same pH conditions for 8 µM α_3C (D), 10 µM α_3W (E), and 15 µM PRM (F). The luminescence spectra of 8 µM α_3C (G) and 15 µM α_3W (H) were collected upon excitation at 340 nm using Ex slit = 5 nm and Em slit = 15 nm. The luminescence spectra of 100 µM PRM (I) was collected upon excitation at 260 nm using Ex slit = 5 nm and Em slit = 15 nm. Percent change in α_3C (J), α_3W (K), and PRM (L) ProCharTS absorbance (with respect to ProCharTS in water) due to extreme pH was monitored at multiple wavelengths.

α_3C ProCharTS absorbance decreased slightly at pH 1 and more abruptly at pH 13 (Figure 4.1A). Similar trend was observed for α_3W with a moderate decrease in ProCharTS at pH 1 and substantial decrease at pH 13 (Figure 4.1B). At extreme pH, the Lys-Glu sidechain electrostatic interactions would reduce; hence, SS-CT transitions are expected to be significantly lowered. Especially the observed decrease in ProCharTS at longer wavelengths was attributed to reduced SS-CT. α_3C ProCharTS at pH 1 between 275—325 nm overlapped with α_3C ProCharTS at neutral pH (Figure 4.1J). The overlap occurred because the secondary structure at pH 1 was not disrupted significantly and charged amino acid contacts were retained (Figure 4.1D). Alkaline pH perturbed α_3C and α_3W secondary structure significantly causing protein unfolding at pH 13 (Figure 4.1D and 4.1E). An interesting observation was made regarding α_3W ProCharTS at pH 1. The secondary structure content remained similar (Table 4.1), suggesting that most charged amino acid contacts were retained in the presence of 0.1 N HCl. Hence the observed decrease in α_3W ProCharTS (at pH 1) was attributed wholly to neutralization of the Glu COO^- groups at pH 1 (Figure 4.1B). Protonation of Glu sidechain eliminating the charge would reduce SS-CT and PBS-CT involving Glutamate.

α_3C and α_3W ProCharTS luminescence did not change significantly at extreme pH (Figure 4.1G and 4.1H) and the observed changes in luminescence were more difficult to correlate with structural transitions. The difficulties in explaining the effect of pH on ProCharTS luminescence reiterated that the phenomenon of charge transfer luminescence was not well understood.

PRM ProCharTS decreased at pH 1 and remained the same at pH 13 (Figure 4.1C). Similarity in PRM secondary structure at pH 7 and pH 1 was elucidated by the overlap between CD spectra (Figure 4.1F) and equivalent secondary structure content (Table 4.1). The decrease in PRM luminescence at pH 1 (Figure 4.1I) correlated with a substantial reduction in PRM absorbance, but the reason for this decrease compared to neutral pH was unclear. PRM absorbance at neutral pH and pH 13 overlapped between 250—290 nm and PRM luminescence at pH 7 and 13 also had similar intensity. The predicted pI of salmon PRM is 13.3 and Arg sidechains were possibly not neutralized enough to change the PRM absorbance and luminescence at pH 13. Because of detector saturation in CD spectra of PRM at pH 13, information about secondary structure parameters could not be obtained. However, visual inspection of the similarity of ellipticity at 222 nm ruled out gain in helicity and similarity of ellipticity till ~212 nm hinted at the absence of β-sheet at alkaline pH.

SAMPLE	α-helix	β-sheet	Random coil
α_3C in H_2O	92.86	0.05	7.09
α_3C in 0.1 N HCl	94.96	0.04	5.00
α_3C in 0.1 N NaOH	22.00	31.32	46.68
α_3W in H_2O	95.08	0.04	4.88
α_3W in 0.1 N HCl	94.91	0.04	5.05
α_3W in 0.1 N NaOH	61.20	2.12	36.68
PRM in H_2O	7.35	14.44	78.21
PRM in 0.1 N HCl	4.24	14.24	81.52

Table 4.1: Monitoring the effect of extreme pH on α_3C, α_3W, and PRM by secondary structure analysis. The helix, sheet, and random coil parameters obtained after CD data analysis by K2D3 program.

4.1.2 The effect of salt on α_3C, α_3W, and PRM ProCharTS

α_3C, α_3W, and PRM ProCharTS absorbance, luminescence and secondary structure was measured in the presence of salt. The presence of salt counterions neutralized charged amino acid sidechains and the ensuing effects on ProCharTS was monitored.

α_3C, α_3W, and PRM ProCharTS absorbance decreased in the presence of 250 mM NaCl (Figure 4.2A—C). α_3C absorbance decreased moderately (~20 %), whereas α_3W and PRM absorbance decreased significantly. α_3W absorbance decreased by more than 50 % in the 350—800 nm region (Figure 4.2K). The decrease in absorbance was observed without any significant change in protein secondary structure (Figure 4.2D—F). While α_3C and α_3W retained their helical content in the presence of 250 mM NaCl, PRM remained disordered (Table 4.2). The observed decrease in α_3C, α_3W, and PRM absorbance was attributed to the neutralization of charged amino acid sidechains by salt counterions. Thus the importance of protein charge in the observed novel absorbance was ascertained by monitoring the decrease in ProCharTS absorbance upon screening charged amino acid sidechains by Na^+ and Cl^- ions. α_3C, α_3W, and PRM ProCharTS luminescence did not change significantly in the presence of salt with α_3C and α_3W showing a slight decrease (Figure 4.2G—I).

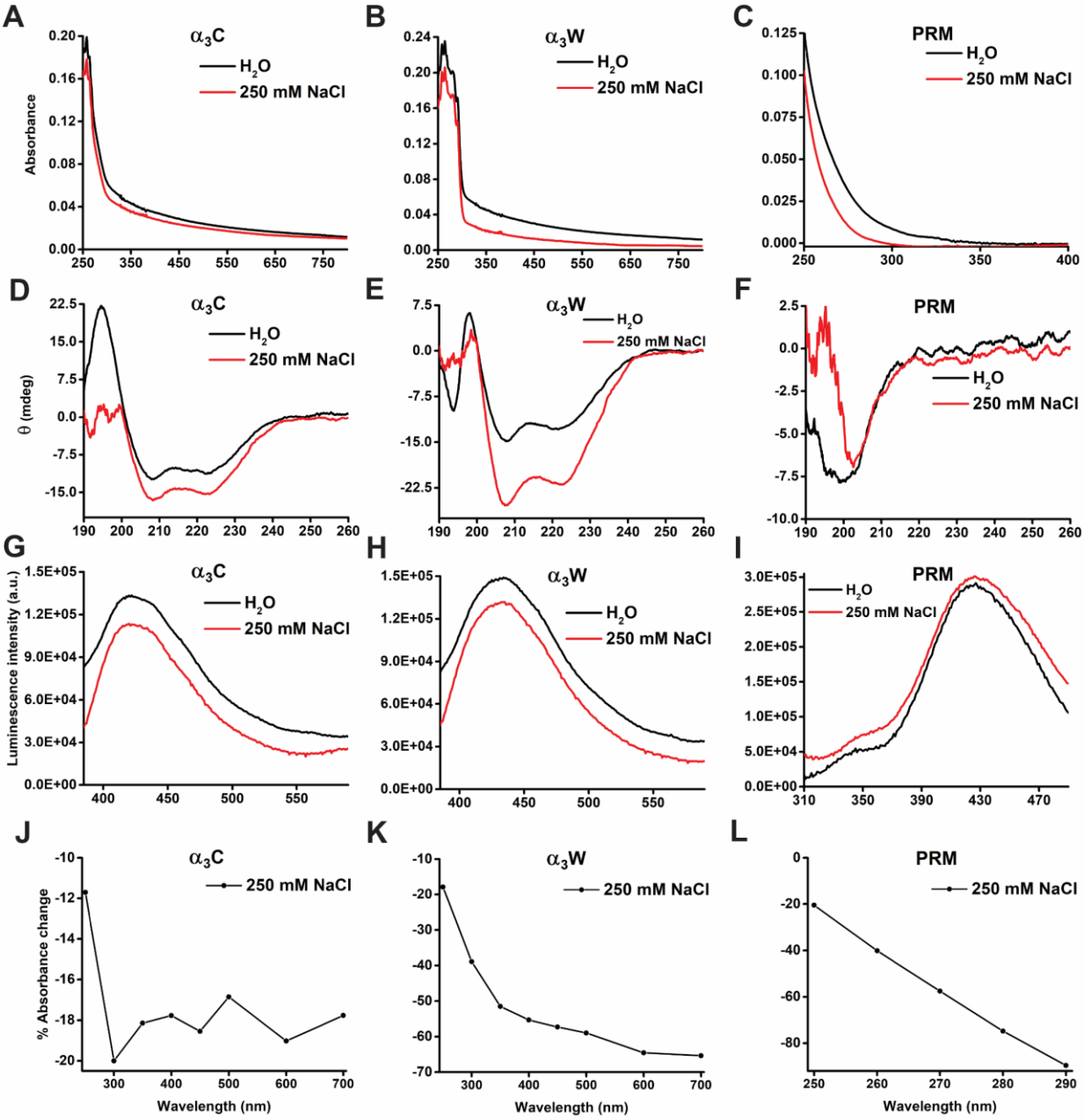


Figure 4.2: Monitoring the effect of salt on α3C, α3W, PRM ProCharTS and secondary structure. The absorbance spectra of 25 µM α3C (A), 25 µM α3W (B), and 200 µM PRM (C) was measured in deionized water and 250 mM NaCl. The CD spectra were recorded for 8 µM α3C (D), 8 µM α3W (E), and 15 µM PRM (F) in deionized water and 250 mM NaCl. The luminescence spectra of 25 µM α3C (G) and 25 µM α3W (H) were collected upon excitation at 340 nm using Ex slit = 5 nm and Em slit = 15 nm. The luminescence spectra of 100 µM PRM (I) was collected upon excitation at 260 nm using Ex slit = 5 nm and Em slit = 15 nm. Percent change in α3C (J), α3W (K), and PRM (L) ProCharTS absorbance (with respect to ProCharTS in water) in the presence of salt was monitored at multiple wavelengths.

α_3C, α_3W, and PRM secondary structure did not change in the presence of 250 mM NaCl (Table 4.2). α_3C and α_3W retained their helicity and PRM remained a disordered protein.

SAMPLE	α-helix	β-sheet	Random coil
α_3C in H_2O	92.91	0.04	7.05
α_3C + 250 mM NaCl	95.09	0.04	4.87
α_3W in H_2O	94.96	0.05	5.00
α_3W + 250 mM NaCl	95.20	0.05	4.75
PRM in H_2O	3.44	13.51	83.05
PRM + 250 mM NaCl	7.35	14.44	78.21

Table 4.2: Monitoring the effect of salt on α_3C, α_3W, and PRM by secondary structure analysis. The helix, sheet, and random coil parameters obtained after CD data analysis by K2D3 program.

4.1.3 The effect of moderate pH changes on α_3C, α_3W, and PRM ProCharTS

The influence of moderate changes in pH on α_3C, α_3W, and PRM ProCharTS absorbance, luminescence, and protein secondary structure was studied. Structural perturbations induced by moderate pH change would be lesser than extreme pH. Additionally, the charged amino acid sidechains would be neutralized to a lesser extent. The effects of these moderate pH changes on protein charge transfer would help get a better understanding of the observed phenomenon.

A prompt decrease in α_3C and α_3W ProCharTS absorbance was observed at pH 3 and 5 compared to neutral pH (Figure 4.3A and 4.3B). The observed decrease was greater at longer wavelengths suggesting that SS-CT reduced due to protonation of Glu sidechains with pH lowering. α_3C and α_3W ProCharTS luminescence decreased substantially at pH 3 and 5 (Figure 4.3G and 4.3H). The observed changes in α_3C and α_3W ProCharTS were attributed to the moderate neutralization of Glu sidechains as changes in secondary structure were absent (Figure 4.3D and 4.3E). Absence of change in secondary structure content was ascertained by analysing α_3C and α_3W CD spectroscopy data (Table 4.3). Due to precipitation of α_3C and α_3W at pH 9 and 11, the effect of increasing pH could not be characterized.

PRM ProCharTS absorbance was the lowest at pH 3 and increased at each successive pH monitored (Figure 4.3C). PRM absorbance at pH 3 decreased more than 40 % (Figure 4.3L) while the absorbance at pH 11 increased marginally in the 260—300 nm region compared to neutral pH. PRM ProCharTS luminescence was affected by pH significantly (Figure 4.3I). At pH 3 and 5, PRM luminescence intensity was low. PRM luminescence increased at pH 9 and 11 and almost overlapped with each other. PRM absorbance and luminescence increased with pH, but differences in absorbance and luminescence increase were observed. For instance, the overlap of PRM absorbance at pH 5 and 9 was absent in PRM luminescence.

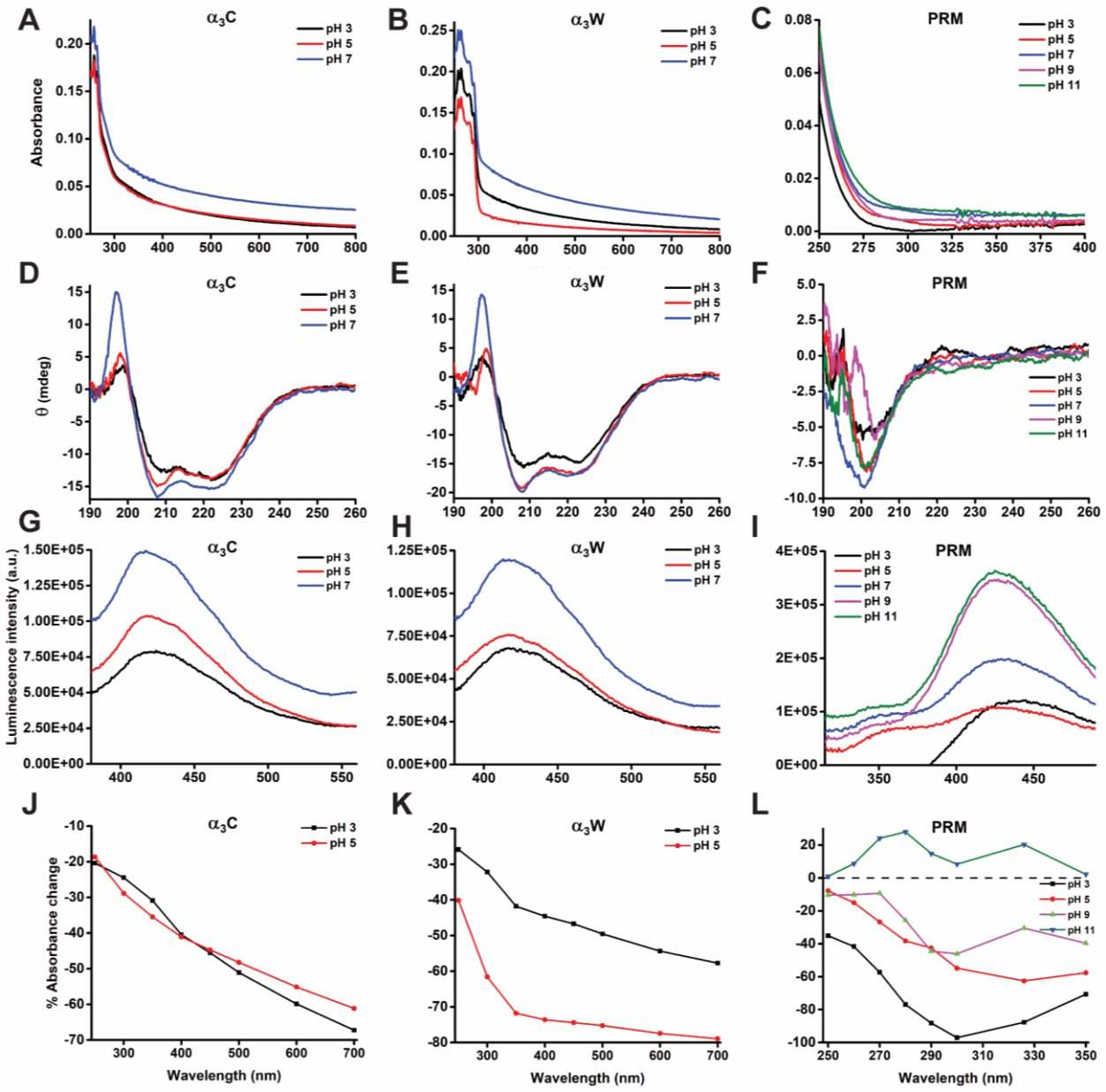


Figure 4.3: Monitoring the effect of moderate pH change on α_3C, α_3W, PRM ProCharTS and secondary structure. The absorbance spectra of 25 µM α_3C (A), 25 µM α_3W (B), and 100 µM PRM (C) was measured at buffered pH 3—11. The CD spectra were recorded in the same pH conditions for 10 µM α_3C (D), 10 µM α_3W (E), and 15 µM PRM (F). The luminescence spectra of 25 µM α_3C (G) and 25 µM α_3W (H) were collected upon excitation at 340 nm using Ex slit = 5 nm and Em slit = 15 nm. The luminescence spectra of 100 µM PRM (I) was collected upon excitation at 260 nm using Ex slit = 5 nm and Em slit = 15 nm. The buffered solutions were prepared at pH 3 (20 mM citric acid), pH 5 (20 mM Na acetate), pH 7 (20 mM NaH_2PO_4), pH 9 (20 mM Tris) and pH 11 (20 mM Na_2HPO_4). Percent change in α_3C (J), α_3W (K), and PRM (L) ProCharTS absorbance upon changing pH (compared to ProCharTS at pH 7) was monitored at multiple wavelengths.

PRM ellipticity at 222 nm was marginally lower at pH 11 (Figure 4.3F) compared to pH 3 (decrease from 0.53 to -1.13 mdeg) however PRM secondary structure analysis revealed the

absence of any detectable structural changes with pH (Table 4.3). The effect of pH on PRM ProCharTS illustrated the difficulty in correlating changes in ProCharTS with solvent parameters.

SAMPLE	α-helix	β-sheet	Random coil
α_3C pH 3	95.16	0.03	4.81
α_3C pH 5	94.95	0.04	5.01
α_3C pH 7	95.24	0.04	4.74
α_3W pH 3	95.09	0.04	7.05
α_3W pH 5	95.07	0.05	7.06
α_3W pH 7	95.07	0.05	7.06
PRM pH 3	4.46	14.14	81.70
PRM pH 5	5.30	12.64	82.06
PRM pH 7	4.59	13.03	82.38
PRM pH 9	7.51	13.53	78.96
PRM pH 11	3.05	12.69	84.26

Table 4.3: Monitoring the effect of moderate pH change on α_3C, α_3W, and PRM by secondary structure analysis. The helix, sheet, and random coil parameters obtained after CD data analysis by K2D3 program.

4.2 Monitoring the effect of ProCharTS luminescence on Trp fluorescence

There is significant overlap between absorbance and luminescence spectra of protein charge transfer states with the Trp absorbance and fluorescence spectra. To understand the contribution of CT states to Trp fluorescence, the effect of protein α_3C on NATA fluorescence intensity decay was monitored. The contribution of ProCharTS luminescence to Trp fluorescence in the single Trp containing protein α_3W was also studied.

4.2.1 The effect of CT states on NATA fluorescence decay

NATA is the structural analogue of Trp in which the N- and C-terminus are capped with acetyl and amide groups respectively. The fluorescence intensity decay of NATA in aqueous solution fits to a single exponential. To assess the effect of protein charge transfer on Trp fluorescence, 20 µM NATA was mixed with increasing concentrations (5—15 µM) of α_3C. This approach was a simplified way of placing charge transfer states in the proximity of Trp to investigate the possible influence of ProCharTS luminescence on Trp fluorescence.

The IRF, fitted, and measured fluorescence intensity decays of NATA in the presence of α_3C are shown in Figure 4.4A and 4.4B. Data fitting was done to a single or sum of two exponentials. Moderate amount of protein (5 µM α_3C) was sufficient to observe the deviation from single exponential fitting as evidenced by the high values of reduced χ^2 and poor residual distribution in Figure 4.4C compared to Figure 4.4D. The reduced χ^2 values corresponding to single exponential fits were much more than 1.2, and the residual distribution was not random.

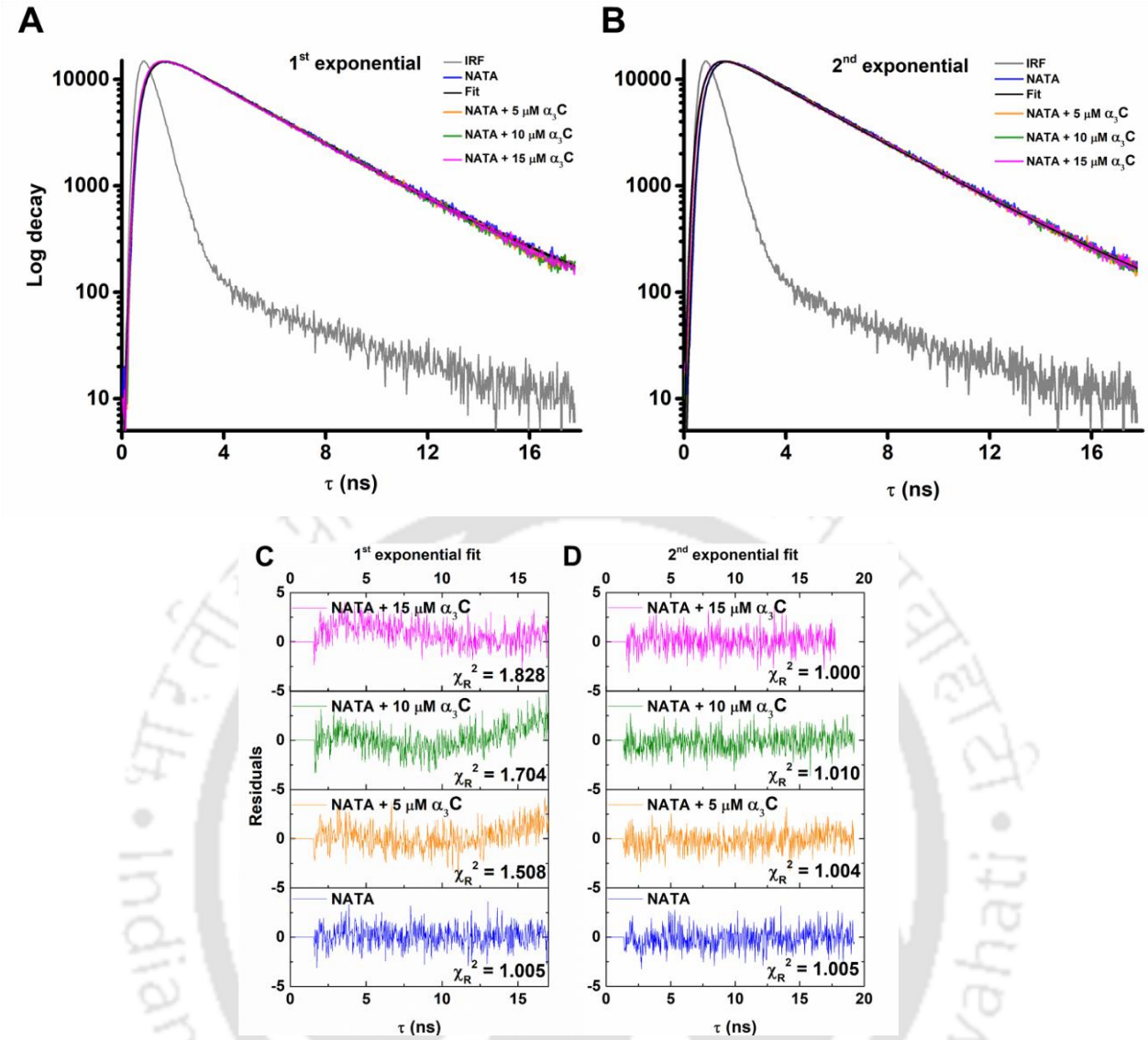


Figure 4.4: Effect of ProCharTS luminescence on NATA fluorescence intensity decay analyzed by single (A) and sum of two exponentials (B) model. Residuals corresponding to single (C) and sum of two (D) exponential fits. 20 μM NATA fluorescence intensity decays were collected using a 295 nm excitation source and a 340 ± 20 nm emission bandpass filter with a TCSPC instrument in the presence of 0—15 μM α_3C. The obtained decays were fit to a single or sum of two exponentials model. The residuals and corresponding χ_R^2 values are reported. All samples were prepared in deionized water.

The contribution of protein charge transfer was uncovered by NATA lifetime decay analysis (Table 4.4). A short component of ~0.9 ns with a small amplitude (0.02—0.04) was detected and attributed to α_3C ProCharTS luminescence. The presence of CT luminescence deviated NATA fluorescence intensity decay from single to sum of two exponentials. Lifetime of the long component contributed by Trp was not quenched in the presence of α_3C. Extremely small amplitude of the ProCharTS luminescence component highlighted that Trp fluorescence was much more intense than CT luminescence and hinted as at why only few reports of this phenomenon have been made till date.

Sample	τ_1 (ns)	α_1	τ_M (ns)	χ_R^2
NATA 20 µM	3.16 {0.016}	1.00	3.16 {0.016}	1.01
NATA + 5 µM α_3C	3.11 {0.018}	1.00	3.11 {0.018}	1.51
NATA + 10 µM α_3C	3.06 {0.003}	1.00	3.06 {0.003}	1.88
NATA + 15 µM α_3C	3.13 {0.008}	1.00	3.13 {0.008}	1.83

Sample	τ_1 (ns)	α_1	τ_2 (ns)	α_2	τ_M (ns)	χ_R^2
NATA 20 µM	3.16 {0.016}	1.00			3.16 {0.016}	1.01
NATA + 5 µM α_3C	3.18 {0.013}	0.98	0.90 {0.209}	0.02	3.12 {0.010}	1.00
NATA + 10 µM α_3C	3.18 {0.015}	0.97	0.82 {0.074}	0.03	3.10 {0.018}	1.01
NATA + 15 µM α_3C	3.16 {0.015}	0.96	0.96 {0.105}	0.04	3.09 {0.002}	1.00

Table 4.4: Effect of ProCharTS luminescence on NATA fluorescence intensity decay. 20 µM NATA fluorescence intensity decays were collected using a 295 nm excitation source and a 340 ± 20 nm emission bandpass filter with a TCSPC instrument in the presence of 0—15 µM α_3C. The obtained decays were fit to a single or sum of two exponentials model. The lifetime (τ_i) and amplitude (α_i) values obtained after data fitting were tabulated and the standard deviation values were reported in parenthesis. The reported τ values are average of three independent measurements, while α values were taken from the best fitted data.

Thus in the absence of protein, NATA fluorescence fit to a single exponential as expected. However, in the presence of α_3C CT states, the fluorescence decay deviated to a sum of two exponentials (Figure 4.5).

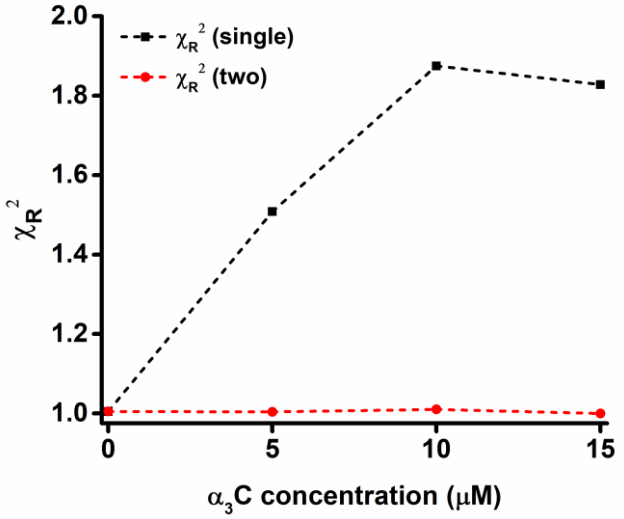


Figure 4.5: Reduced χ^2 analysis corresponding to NATA fluorescence intensity decay in the presence of α_3C. The deviation of NATA fluorescence intensity decay from single to sum of two exponential fits in the presence of 5—15 µM α_3C was shown by the increase in χ_R^2 of single exponential fits. The samples were prepared in deionized water.

Additionally, fluorescence lifetime distribution analysis was done for NATA fluorescence intensity decays in the presence of α_3C using MEM. The amplitude values obtained were normalized to 1 and plotted. A short component originating from ProCharTS was resolved (Figure 4.6A). Smaller amplitude compared to Trp fluorescence reiterated that ProCharTS luminescence had much lower intensity. χ^2 values below 1.2 and randomly distributed residuals illustrated the goodness of fits (Figure 4.6B).

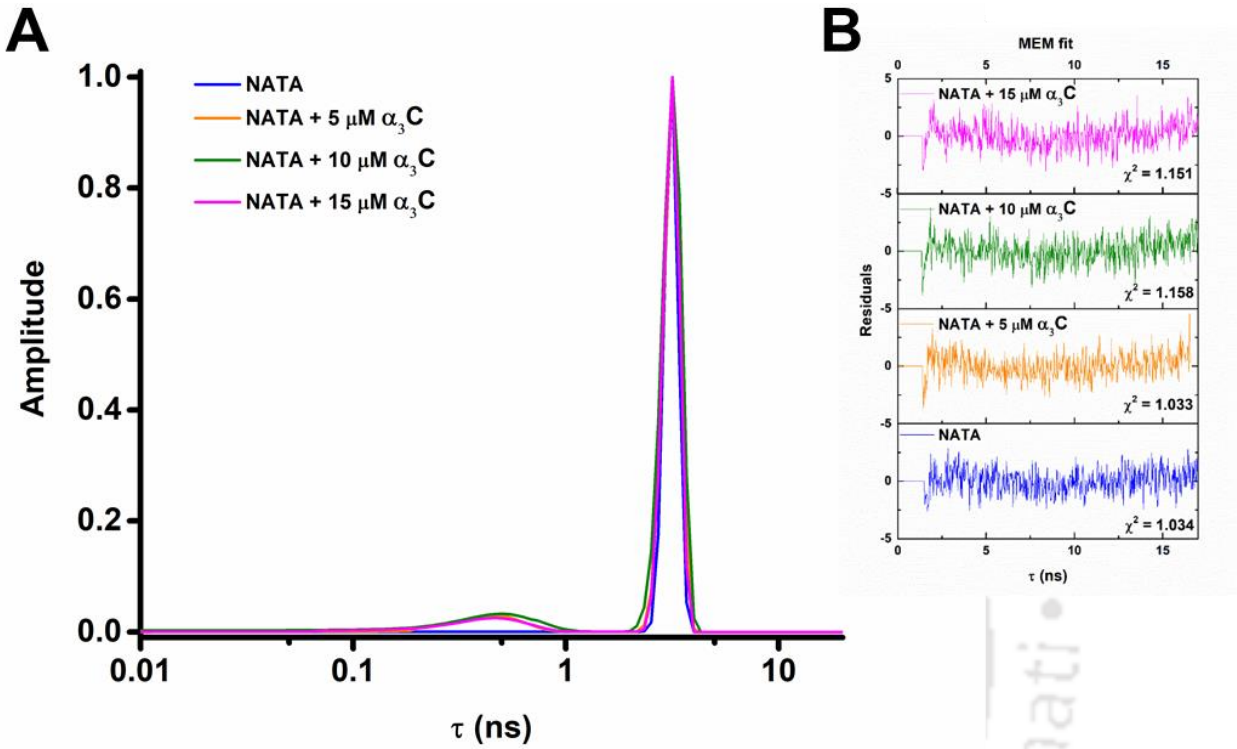


Figure 4.6: (A) Effect of ProCharTS luminescence on NATA fluorescence intensity decay. (B) Residuals corresponding to MEM fits. 20 μM NATA fluorescence intensity decays were collected using a 295 nm excitation source and a 340 ± 20 nm emission bandpass filter with a TCSPC instrument in the presence of 0—15 μM α_3C. The obtained decays were fit by maximum entropy method to obtain the lifetime distribution, the residuals, and corresponding χ^2 values.

The contribution of ProCharTS luminescence to NATA fluorescence was deciphered only by time-resolved method. NATA steady-state fluorescence intensity in the presence of α_3C did not change significantly and the steady-state emission obtained from NATA and NATA + α_3C mixtures upon excitation at 295 nm remained similar in shape and intensity (Figure 4.7). This is consistent with the mean lifetime remaining almost constant with addition of α_3C (Table 4.4).

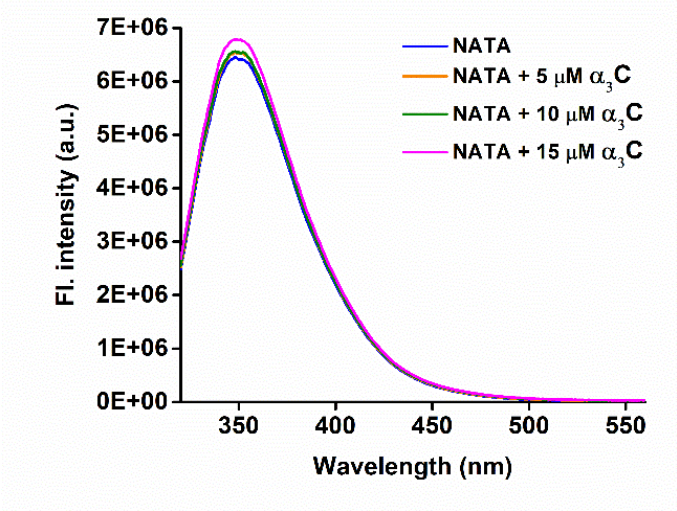


Figure 4.7: Monitoring the effect of ProCharTS luminescence on NATA steady-state fluorescence. 20 μM NATA fluorescence was collected in the presence of 0—15 μM α_3C upon excitation at 295 nm in the 320—560 nm range using Ex slit = 2 nm and Em slit = 15 nm. The samples were prepared in deionized water.

NATA steady-state anisotropy increased from 0.0058 to 0.0091 in the presence of 15 μM α_3C (Figure 4.8). NATA r_{ss} was low in the absence of protein as expected for a small fluorophore which can rotate freely. The marginal increase of NATA r_{ss} in NATA + α_3C mixtures, suggested non-specific interaction of a small fraction of NATA with protein molecules slowing the rotational diffusion of the fluorophore. Additionally, anisotropy increase suggested the proximity of at least some CT states to Trp chromophore in NATA.

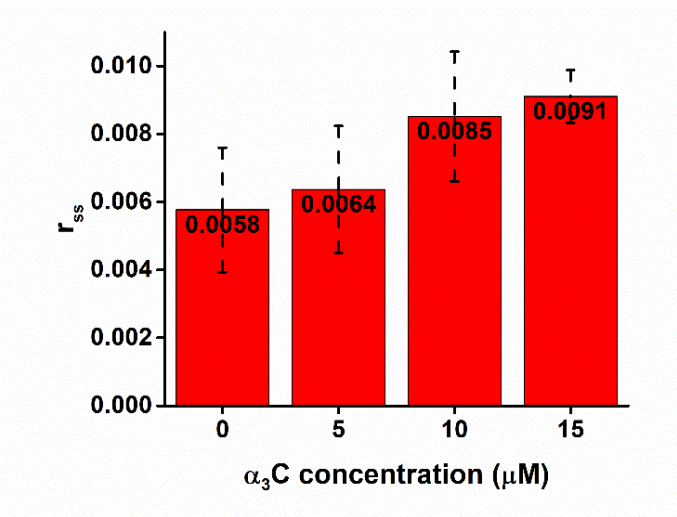


Figure 4.8: Steady-state anisotropy of 20 μM NATA in the presence of 0—15 μM α_3C. The non-specific interaction of NATA with α_3C was checked by NATA steady-state anisotropy upon excitation at 295 nm and collecting the emission at 350 nm with Ex slit = 1 nm and Em slit = 5 nm. The samples were prepared in deionized water.

4.2.2 The effect of CT states on Trp fluorescence in single Trp containing protein α_3W

In the protein α_3W, Trp is in the proximity of charged amino acid sidechains making it a good model system to study the effect of ProCharTS luminescence on Trp fluorescence. To identify the contribution from α_3W CT states, luminescence intensity decays were collected from the proteins α_3C and α_3W upon excitation at 295 nm. A comparison of the lifetime analysis by MEM ascertained the additive effect of ProCharTS to Trp emission.

α_3C and α_3W lifetime distribution analysis were performed after luminescence intensity decays were collected upon excitation at 295 nm. α_3C luminescence originates solely from ProCharTS and comprised of three components (Figure 4.9A) with τ_3 having the smallest amplitude whereas τ_1 had maximum amplitude in agreement with discrete analysis (Table 4.5). Differences compared to discrete analysis arose in the MEM fit of α_3W where a short component originating from ProCharTS was resolved and appeared as a broad feature with no specific peaks implying that in α_3W, both Trp fluorescence and ProCharTS contributed to luminescence decay (Figure 4.9A). The huge difference in amplitude between Trp fluorescence and ProCharTS luminescence in α_3W reasserted that Trp fluorescence was much more intense than ProCharTS luminescence. The goodness of data fitting was assessed by monitoring the random distribution of residuals and the acceptable χ^2. The fit quality was better for α_3W compared to α_3C: α_3C χ^2 was higher but the residuals were randomly distributed (Figure 4.9B). The fit quality of α_3C was poorer compared to α_3W because longer decay acquisition time in α_3C resulted in higher background counts.

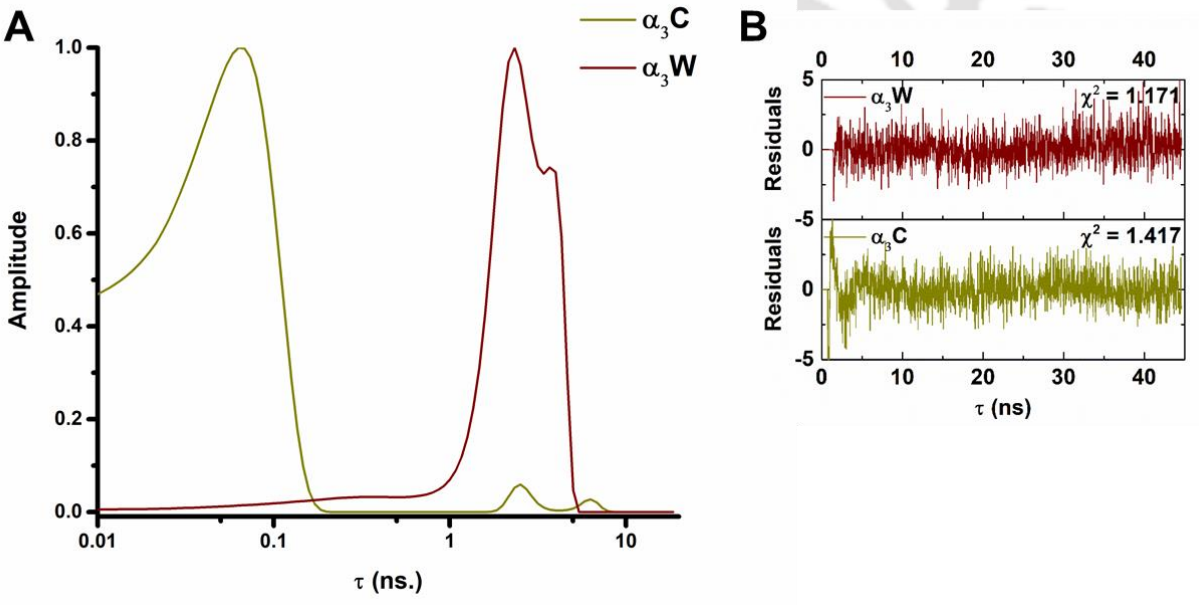


Figure 4.9: (A) α_3C and α_3W luminescence intensity decay analysis by Maximum Entropy Method. (B) Residuals for the corresponding analysis. 15 µM α_3C and α_3W samples were prepared in deionized water and luminescence decays were collected upon excitation at 295 nm using a 340 ± 20 nm emission bandpass filter in a TCSPC instrument. For the plotted data, the amplitude values were normalized to 1. The residuals and corresponding χ^2 values were reported.

The IRF, fitted, and measured intensity decays of α_3C and α_3W upon excitation at 295 nm are shown in Figure 4.10A and 4.10B. α_3C and α_3W intensity decays were analysed by discrete method using NLLS fitting to obtain individual lifetime components and corresponding amplitudes. α_3C luminescence intensity decay fit to a sum of three exponentials and α_3W fluorescence intensity decay fit to a sum of two exponentials. Reduced χ^2 values below 1.2 and randomness in residual distribution implied that data fit well to assumed models (Figure 4.10C and 4.10D). The contribution of ProCharTS luminescence to Trp fluorescence in α_3W was not detected by discrete analysis of intensity decay data.

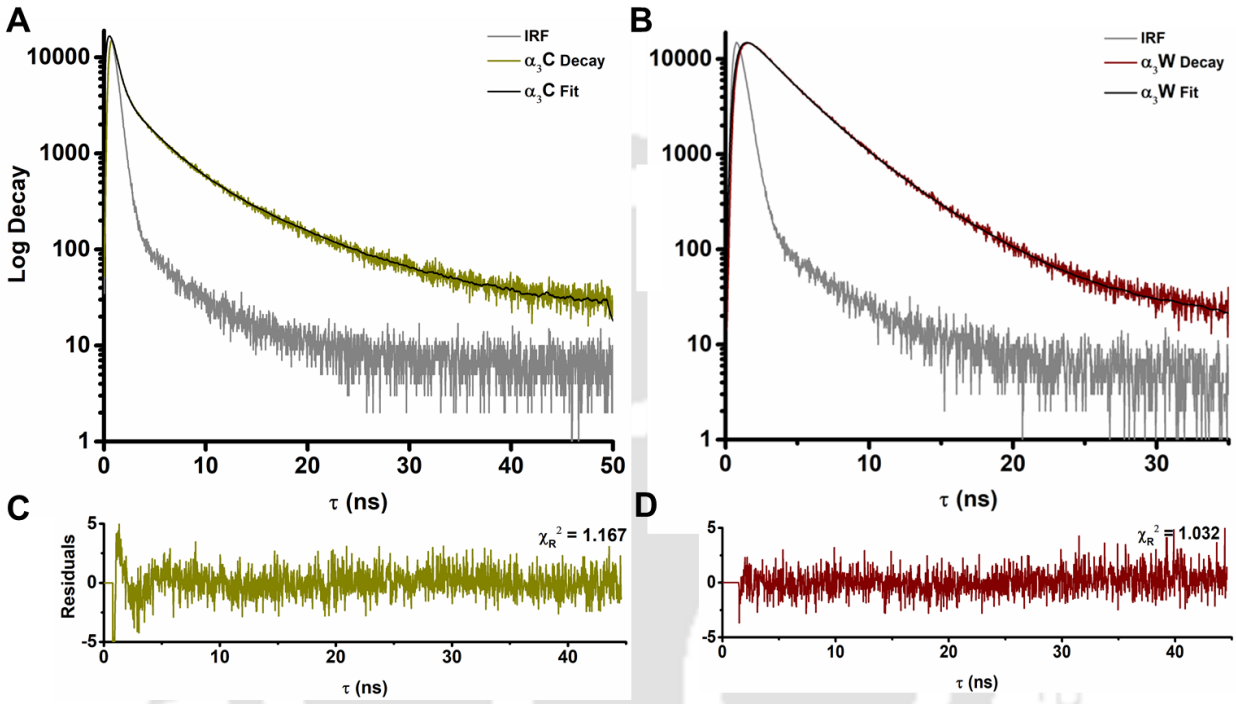


Figure 4.10: α_3C (A) and α_3W (B) luminescence intensity decay analysis by discrete method. (C) and (D) Residuals for the corresponding analysis. 15 µM α_3C and α_3W samples were prepared in deionized water and luminescence decays were collected upon excitation at 295 nm using a 340 ± 20 nm emission bandpass filter in a TCSPC instrument. α_3C fit to a sum of three exponentials and α_3W fit to a sum of two exponentials model. The residuals and corresponding χ_R^2 values were reported.

α_3C mean lifetime was 0.70 ns (Table 4.5). α_3C lifetime had three components: short component having large amplitude and long component with smallest amplitude. This was in agreement with α_3C lifetime distribution analysis by MEM. The short component τ_1 of 0.37 ns in α_3C had an amplitude of 0.897. Two components originating from Trp were detected in α_3W fluorescence decay analysis. α_3W mean lifetime was 2.8 ns implying that Trp fluorescence was not quenched significantly by the CT states present in Trp proximity.

Sample	τ_1 (ns)	α_1	τ_2 (ns)	α_2	τ_3 (ns)	α_3	τ_M (ns)	χ_R^2
α_3W			2.30 {0.116}	0.763	4.29 {0.241}	0.237	2.80 {0.058}	1.030
α_3C	0.37 {0.006}	0.897	2.68 {0.135}	0.085	8.34 {0.435}	0.018	0.70 {0.038}	1.167

Table 4.5: α_3C and α_3W luminescence intensity decay analysis by discrete method. The lifetime (τ_i) and amplitude (α_i) values obtained after fitting α_3C luminescence decay to a sum of three exponentials and α_3W decay to a sum of two exponentials model was tabulated. The standard deviation values were reported in parenthesis. The reported τ values are average of three independent measurements, while α values were taken from the best fitted data.

4.2.3 Assessing the influence of ProCharTS on Trp steady-state fluorescence

The additive influence of ProCharTS luminescence was also assessed by comparing the steady-state fluorescence intensity of α_3W with already published RNase T1 data.[92] RNase T1 was chosen as its emission maximum is extremely blue shifted like α_3W implying that the Trp in these two proteins are buried in the hydrophobic core. RNase T1 emits fluorescence till 400 nm and there is absence of emission at longer wavelengths (Figure 4.11). Compared to RNase T1, α_3W shows luminescence beyond 400 nm extending the tail till 450 nm. This long tail possibly originates due to CT luminescence and is curtailed in RNase T1 due to comparatively lower charge content (16 % charge in RNase T1 compared to 54 % charge in α_3W).

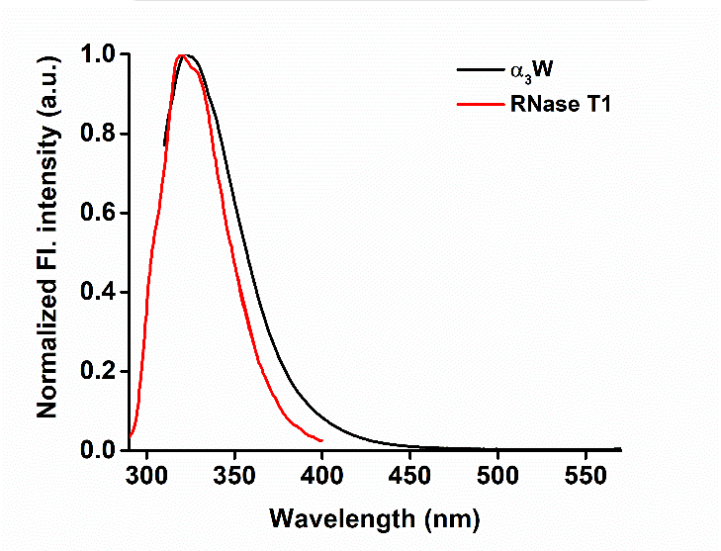


Figure 4.11: Comparison of α_3W and RNase T1 steady-state fluorescence. The steady-state emission of α_3W was collected upon excitation at 295 nm, the fluorescence normalized and compared to the emission from RNase T1 (RNase T1 fluorescence data published by Keifhaber *et al*, 1992).[92]

4.3 Conclusions

1. Solvent parameters such as pH and salt influence ProCharTS absorbance and luminescence by their electrostatic effects on charged amino acid sidechains. The neutralization of Glu side chains at pH 1 significantly reduced α_3W ProCharTS at acidic pH.
2. α_3C and α_3W secondary structure disruption at pH 13 was accompanied by significant decrease in ProCharTS due to reduction in charged amino acid contacts upon protein unfolding.
3. PRM ProCharTS increased with pH without any structure gain. PRM remained intrinsically disordered with changing pH.
4. ProCharTS absorbance decrease in the presence of NaCl was attributed to electrostatic screening of charged amino acid sidechains decreasing the extent of protein charge transfer.
5. The effects of change in solvent parameters on ProCharTS luminescence were less straightforward and highlighted that the phenomenon of luminescence is not well understood.
6. The additive effect of ProCharTS luminescence on Trp emission was observed by analysis of NATA fluorescence intensity decay. The deviation of NATA fluorescence decay fit from single to a sum of two exponentials in the presence of α_3C CT states was elucidated.
7. MEM analysis of α_3C and α_3W luminescence decays was performed. The contribution of both ProCharTS luminescence and Trp fluorescence was observed in the lifetime distribution analysis of α_3W by the maximum entropy method.

CHAPTER 5

THE APPLICATION OF PROTEIN CHARGE TRANSFER SPECTRA TO MONITOR PROTEIN UNFOLDING

5.1 The application of protein charge transfer spectra to monitor protein unfolding

Chemical denaturant-induced HSA, α_3W, and α_3C unfolding was characterized using ProCharTS and conventional methods such as CD spectroscopy, α_3W and HSA Trp fluorescence, α_3C-Dansyl and HSA-Dansyl fluorescence. HSA, α_3W, and α_3C are abundant in charged amino acid sidechains in proximity. Protein unfolding would be accompanied by significant disruption of these contacts, and it was investigated whether unfolding would result in decreased ProCharTS absorbance and luminescence. Differences compared to the characterization of protein unfolding by conventional methods could arise. Hence parameters obtained by nonlinear regression analysis after monitoring protein unfolding with ProCharTS were compared to conventional methods. Inbuilt functions in Origin software were used to perform nonlinear regression and the equations used are given in the appendix.

5.1.1 The application of protein charge transfer absorbance spectra to monitor HSA unfolding

HSA unfolding was performed by the chemical denaturant GdnHCl and monitored using HSA ProCharTS at multiple wavelengths and far UV CD spectroscopy. HSA ProCharTS decreased substantially upon unfolding (Figure 5.1A), with $Abs_{340\ nm}$ decreasing from 0.079 to 0.027 in the presence of 6 M GdnHCl (Figure 5.1B). A significant decrease in HSA ProCharTS was seen even in the presence of 1 M GdnHCl, and the absorbance decreased to a baseline value at 3 M GdnHCl concentration. The decrease in ProCharTS was attributed to the loss of charged amino acid sidechain contacts due to protein unfolding. HSA unfolding was concomitantly monitored by observing $MRE_{222\ nm}$ increase upon chemical denaturation (Figure 5.1D). The CD spectra of HSA in the absence of denaturant had the characteristic features of an α-helix, and upon unfolding became a predominantly random coil (Figure 5.1C).

ProCharTS was additionally able to detect a molten globule-like state previously reported for HSA observed in the presence of sub-denaturing GdnHCl concentrations.[6] HSA ProCharTS absorbance increased in 0.5 M GdnHCl compared to the sample without denaturant hinting toward the presence of this molten globule state. The number of charged amino acid contacts in the compact molten globule state possibly increased, contributing to increased ProCharTS absorbance. HSA CD spectra in the presence of 0.5 M GdnHCl overlapped with the CD spectra of the sample without denaturant. The $MRE_{222\ nm}$ did not change significantly in the presence of 0.5 M GdnHCl as expected for a molten globule, in agreement with a previous report.[6] Absence of helical content was observed in the presence of 4 M and higher GdnHCl concentration. Nonlinear regression analysis was used to estimate the denaturation midpoint of HSA unfolding. To reduce the numerous disulfide bonds present in HSA and ensure complete protein unfolding, HSA unfolding experiments were performed in the presence of 0.5 mM TCEP.

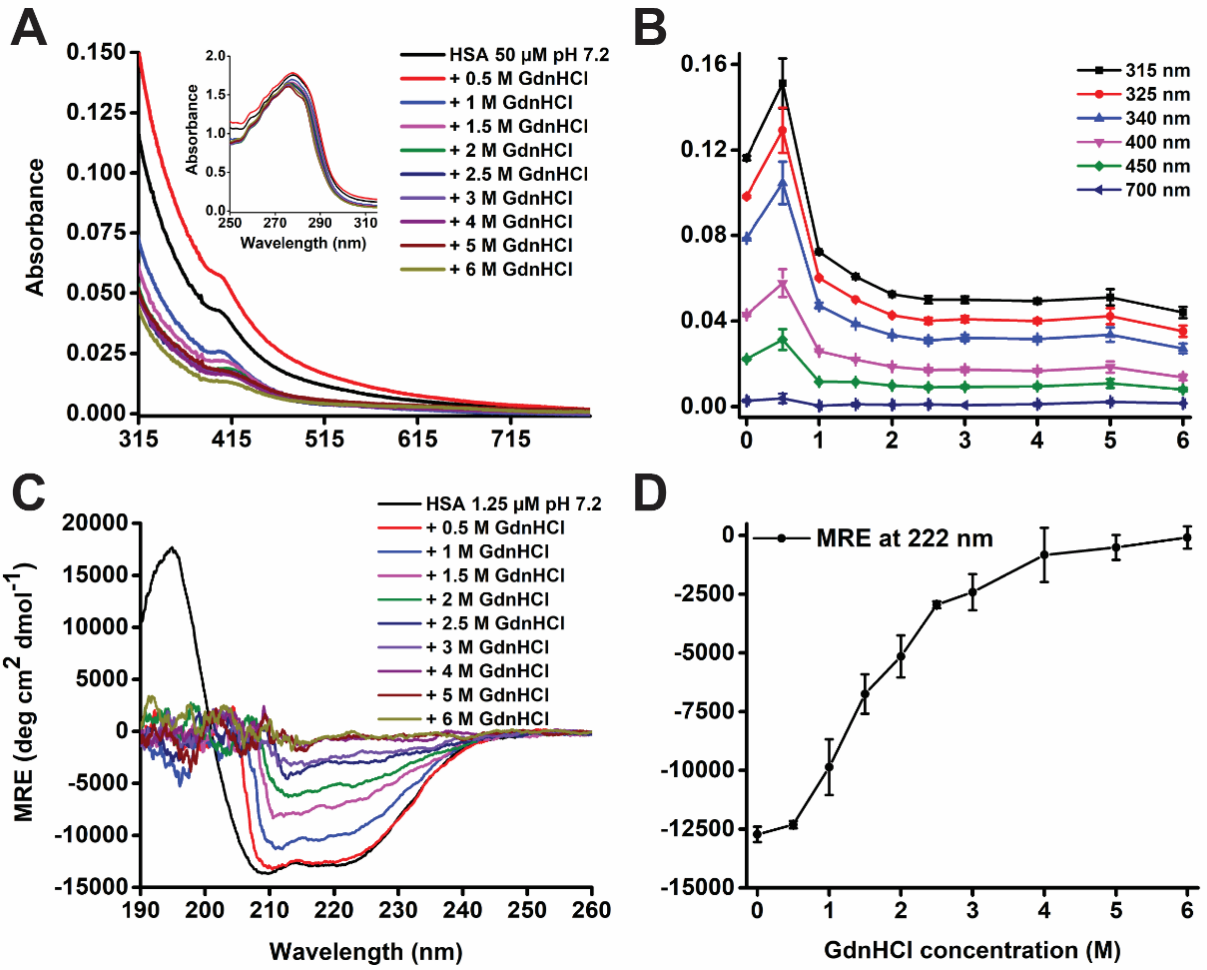


Figure 5.1: Monitoring HSA unfolding by GdnHCl using ProCharTS and CD spectroscopy. 50 µM HSA was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. Absorbance spectra were recorded between 250—800 nm (Figure 5.1A) for three independent samples. The absorbance at multiple wavelengths was monitored upon HSA unfolding and represented as a scatter plot (Figure 5.1B). CD spectra were recorded for samples diluted to 1.25 µM HSA concentration in the 190—260 nm region (Figure 5.1C) and the $MRE_{222\ nm}$ was reported for two independent measurements (Figure 5.1D). Trp absorbs strongly in the 250—315 nm region, with both ProCharTS and aromatic amino acid absorbance expected in this region which is plotted in the inset of Figure 5.1A.

The decrease in $Abs_{315\ nm}$, $Abs_{325\ nm}$, $Abs_{340\ nm}$, and a concomitant increase in $MRE_{222\ nm}$ values (Table 5.1) pointed to the disruption of charged amino acid sidechain contacts and loss of secondary structure upon protein unfolding. The secondary structure parameters were obtained by analysis of HSA CD spectra collected in the presence of increasing GdnHCl concentration. HSA helical content decreased substantially in the presence of 2 M GdnHCl, and complete unfolding was observed at 4 M GdnHCl and beyond. Secondary structure analysis showed an apparent increase in β-sheet content which is not expected in the presence of high GdnHCl concentrations. Based on NMR and radius of gyration analysis, Matsuo *et al* interpreted that denatured proteins do not exist as a β-sheet but as an ensemble of multiple short peptide segments with β-strand

dihedral angles.[93] The molten globule-like state corresponded to increased HSA absorbance values in the presence of 0.5 M GdnHCl. HSA unfolding pathway was thus characterized using absorbance and CD spectroscopy.

GdnHCl concentration (M)	$Abs_{315\ nm}$	STDEV	$Abs_{325\ nm}$	STDEV	$Abs_{340\ nm}$	STDEV	$MRE_{222\ nm}$	STDEV	α-helix	β-sheet	Random coil
0	0.116	0.0010	0.098	0.0005	0.079	0.0001	-12722	330	67.57	9.70	22.73
0.5	0.151	0.0116	0.129	0.0105	0.105	0.0100	-12307	155	68.74	9.47	21.79
1	0.072	0.0004	0.060	0.0005	0.047	0.0013	-9861	1191	67.63	9.35	23.02
1.5	0.061	0.0000	0.050	0.0000	0.039	0.0007	-6747	840	66.96	9.42	23.62
2	0.052	0.0002	0.043	0.0002	0.033	0.0002	-5149	895	45.74	11.61	42.65
2.5	0.050	0.0018	0.040	0.0015	0.031	0.0013	-2949	142	21.69	22.20	56.11
3	0.050	0.0016	0.041	0.0016	0.032	0.0015	-2411	767	12.18	26.05	61.77
4	0.049	0.0012	0.040	0.0012	0.032	0.0011	-828	1152	2.17	30.93	66.90
5	0.051	0.0038	0.042	0.0037	0.034	0.0033	-511	532	3.76	29.42	66.82
6	0.044	0.0026	0.035	0.0026	0.027	0.0023	-86	474	1.21	31.60	67.19

Table 5.1: Monitoring HSA unfolding by GdnHCl using ProCharTS and CD spectroscopy. 50 µM HSA was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ at multiple points during HSA unfolding were collected and tabulated. CD spectra were recorded for samples diluted to 1.25 µM HSA concentration, and the $MRE_{222\ nm}$ was reported. Secondary structure analysis parameters obtained by K2D3 software upon HSA unfolding were also tabulated.

The nonlinear regression analysis of absorbance and MRE values obtained at multiple points in the HSA unfolding pathway were used to obtain denaturation midpoints. At 295 nm, both Trp and ProCharTS are expected to contribute to HSA absorbance making the interpretation of denaturation midpoint complicated. $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ are attributed entirely to ProCharTS absorbance. $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ values fit well to the exponential decay function with R^2 values between 0.983—0.999 (Figure 5.2A—C). The presence of a molten globule-like state was observed in the regression analysis as well. ProCharTS at 315—340 nm in the 0.5 M GdnHCl sample was an outlier deviating from the fit curve. $MRE_{222\ nm}$ fit well to the logistic function with a high R^2 value of 0.999 (Figure 5.2D). The denaturation midpoints obtained highlighted that charged amino acid contacts got disrupted earlier than the loss of secondary structure (0.80 and 0.61 M for $Abs_{325\ nm}$ and $Abs_{340\ nm}$ compared to 1.64 M for $MRE_{222\ nm}$). When HSA absorbance values in the presence of 0.5 M GdnHCl were excluded from nonlinear regression analysis, the denaturation midpoints decreased marginally. Similar midpoints of 0.80 and 0.60 M were estimated by $Abs_{325\ nm}$ and $Abs_{340\ nm}$ analysis upon excluding 0.5 M GdnHCl value.

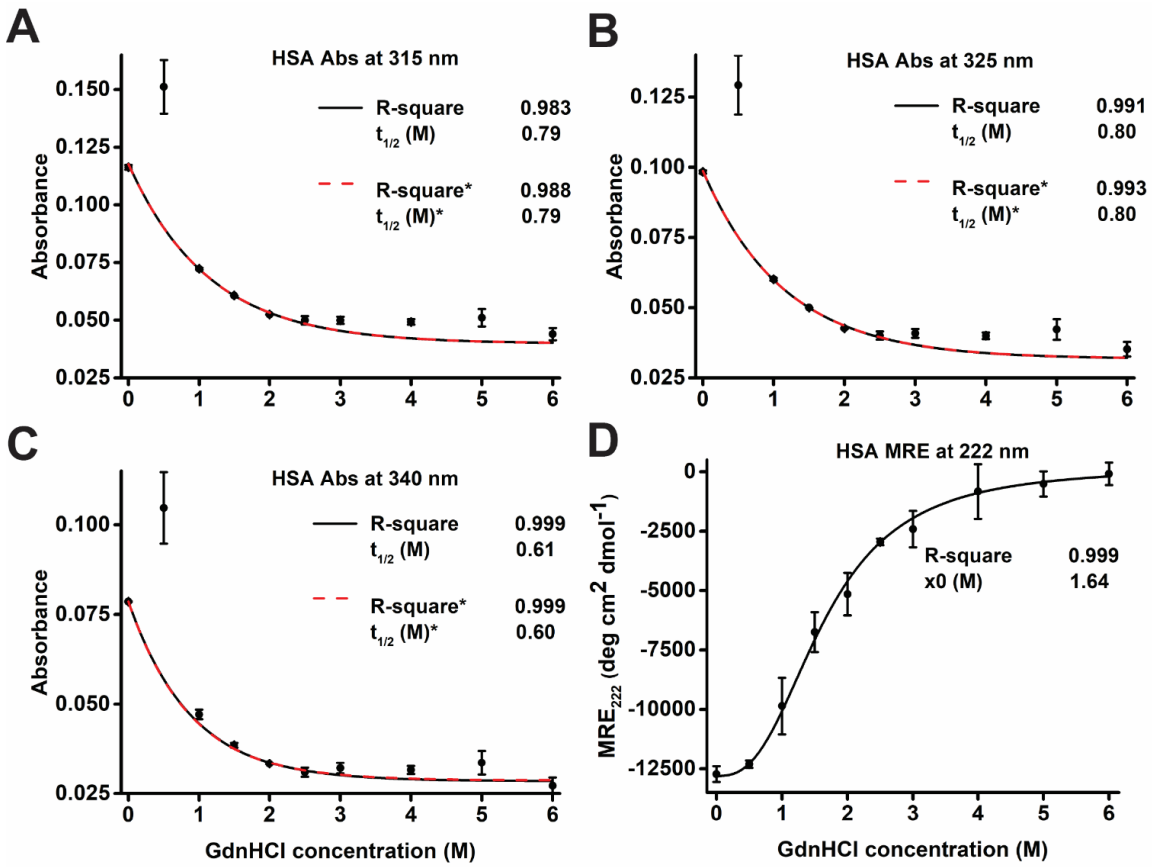


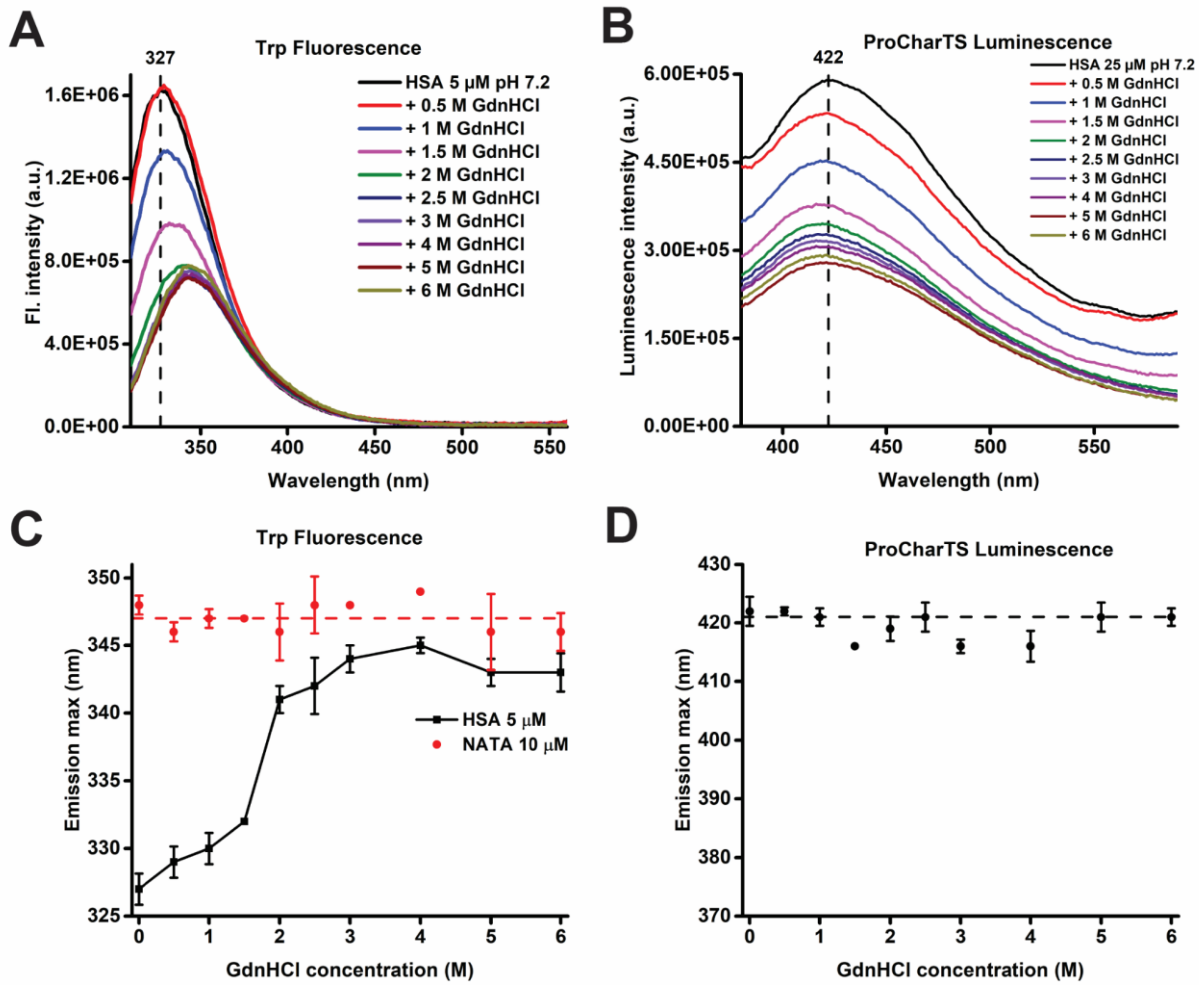
Figure 5.2: Nonlinear regression analysis was used to estimate the denaturation midpoint of HSA unfolding by GdnHCl using ProCharTS and CD spectroscopy. HSA absorbance values at 315 nm (Figure 5.2A), 325 nm (Figure 5.2B), and 340 nm (Figure 5.2C) fit the exponential decay function while MRE data fit the logistic function (Figure 5.2D). Denaturation midpoint values and R^2 values of obtained fits are reported in the figure. The denaturation midpoint and R^2 values obtained upon exclusion of HSA absorbance at 0.5 M GdnHCl from nonlinear regression analysis are reported with an asterisk. The black solid line and red dashed lines represent the fits obtained upon analysis of HSA ProCharTS absorbance with increasing GdnHCl concentration including and excluding the absorbance value at 0.5 M GdnHCl, respectively.

5.1.2 The application of protein charge transfer luminescence spectra to monitor HSA unfolding

Trp fluorescence upon excitation at 295 nm (Figure 5.3A) and ProCharTS luminescence upon excitation at 340 nm (Figure 5.3B) were used to monitor HSA unfolding by GdnHCl. Trp fluorescence intensity (Figure 5.3E) and Trp r_{ss} decreased (Figure 5.3G) while emission maxima red shifted (Figure 5.3C) upon denaturant-induced unfolding. Trp emission maximum shifted from 327 to 343 nm upon protein unfolding. The red shift in emission maxima suggested Trp exposure to solvent and increased solvent relaxation of fluorophore exposed to solvent molecules in the unfolded protein. A decrease in fluorescence intensity upon protein unfolding was observed, and Trp fluorescence reached a baseline value in the presence of 2.5 M GdnHCl. A marginal increase in Trp fluorescence at 0.5 M GdnHCl concentration also hinted at the detection of molten globule-

like intermediate. Trp r_{ss} is sensitive to local as well as global protein dynamics. Both these motions are expected to increase significantly upon protein unfolding. Trp r_{ss} decreased upon protein unfolding due to the higher rotational mobility of Trp fluorophore in the unfolded protein. A slight decrease in fluorescence anisotropy was observed even in the presence of 1 M GdnHCl, and the anisotropy decreased to a baseline value at 3 M GdnHCl. The decrease in HSA Trp fluorescence intensity and red shift in emission maxima were compared to NATA emission. The Trp moiety in NATA is solvent-exposed, and hence negligible changes were observed in the presence of increasing GdnHCl concentration.

HSA Trp fluorescence data was used to correlate with the decrease in ProCharTS luminescence upon protein unfolding (Figure 5.3B). ProCharTS luminescence intensity reduced by ~50 % in the presence of 3 M GdnHCl (Figure 5.3F). Two interesting observations were made: the potential use of a decrease in ProCharTS luminescence to monitor HSA unfolding was uncovered, and luminescence emission maxima were found to be insensitive to protein unfolding (Figure 5.3D). Since the exact nature of the chromophore involved in ProCharTS luminescence is not known conclusively, the study provided an insight into the similarities and differences between the application of Trp fluorescence and the novel ProCharTS luminescence to monitor protein unfolding. Nonlinear regression analysis was used to estimate the denaturation midpoint of HSA unfolding (Figure 5.4).



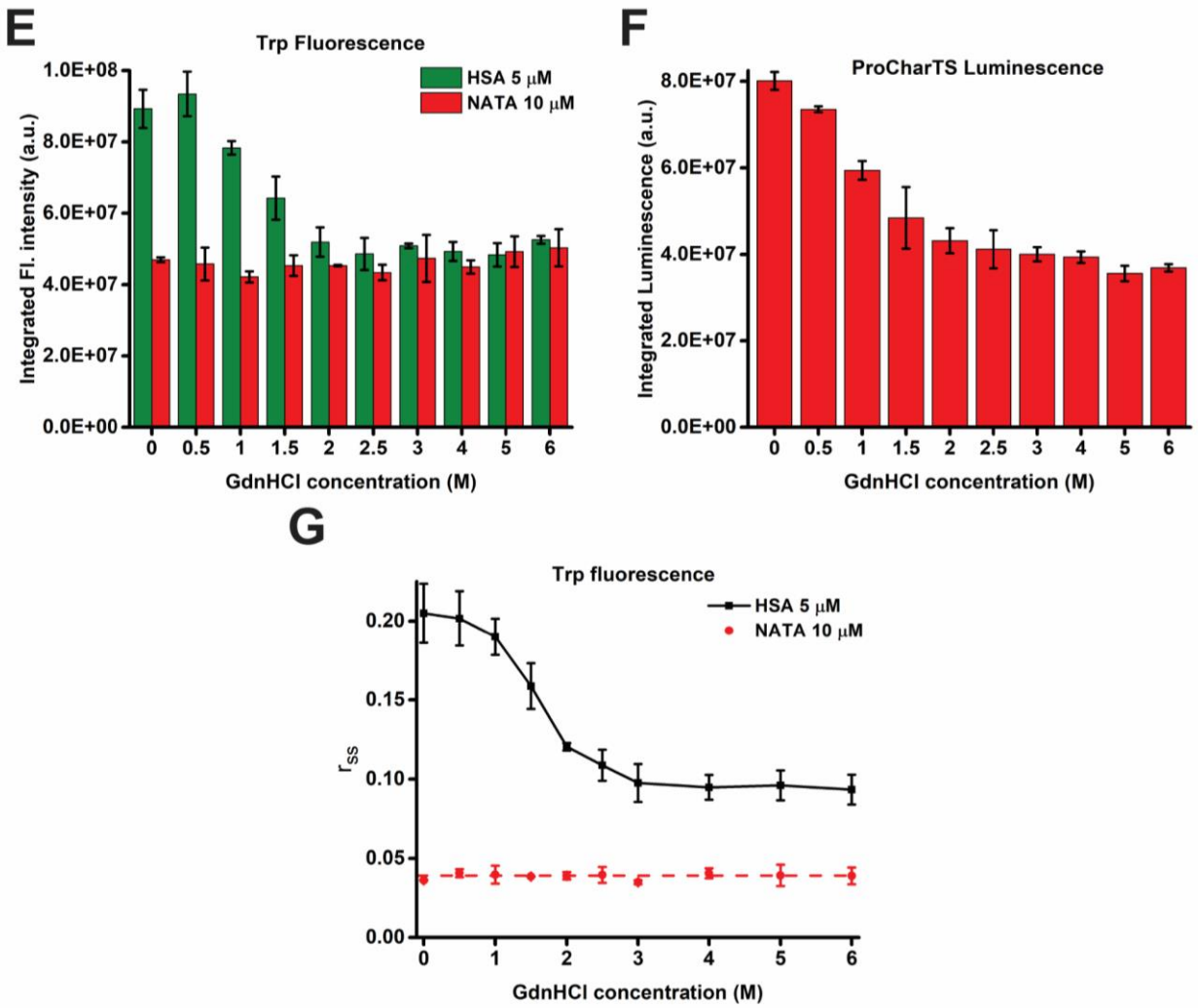


Figure 5.3: Monitoring HSA unfolding by GdnHCl using HSA Trp fluorescence and ProCharTS luminescence. 25 μM HSA was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. Steady-state Trp fluorescence was collected for 5 μM HSA (Figure 5.3A) and 10 μM NATA with Ex. λ = 295 nm and emission in the range of 310—560 nm using Ex. slit = 2 nm and Em. slit = 5 nm. ProCharTS luminescence was collected for 25 μM HSA (Figure 5.3B) with Ex. λ = 340 nm and emission in the range of 365—610 nm using Ex. slit = 2 nm and Em. slit = 15 nm. The Em. λ_{max} for Trp fluorescence (Figure 5.3C) and ProCharTS luminescence (Figure 5.3D) are plotted. Area under the curve of Trp fluorescence (Figure 5.3E) and ProCharTS luminescence (Figure 5.3F) were plotted. HSA Trp steady-state anisotropy was collected with Ex. λ = 295 nm and Em. λ = 330 nm using Ex. slit = 2 nm and Em. slit = 15 nm (Figure 5.3G). The horizontal dashed lines show the absence of change in NATA fluorescence emission max (Figure 5.3C), ProCharTS emission max (Figure 5.3D), and NATA anisotropy (Figure 5.3G) with increasing GdnHCl concentration.

Values for the decrease in Trp fluorescence intensity and red shift in emission maxima were used to monitor HSA unfolding (Table 5.2). NATA fluorescence intensity values were insensitive to increasing denaturant concentration as expected. Additionally, a decrease in ProCharTS luminescence intensity values was also observed upon protein unfolding. Trp fluorescence and

ProCharTS luminescence values at various points in the HSA unfolding pathway are given in the underlying table.

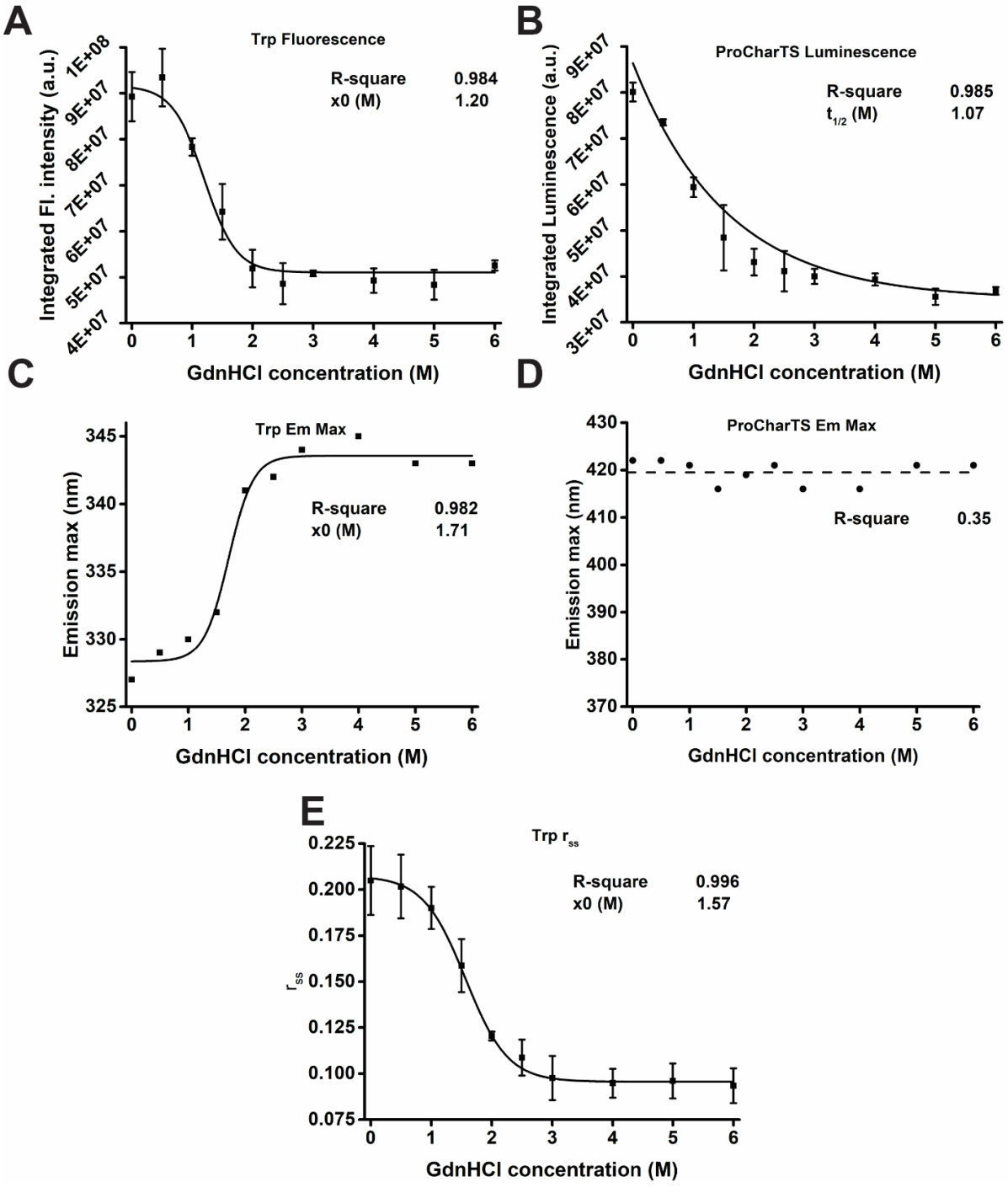
GdnHCl concentration (M)	$Area_{HSA\ 295\ nm}$	STDEV	$Area_{NATA\ 295\ nm}$	STDEV	$Area_{HSA\ 340\ nm}$	STDEV
0	8.93E+07	5.34E+06	4.70E+07	6.85E+05	8.01E+07	2.04E+06
0.5	9.34E+07	6.26E+06	4.58E+07	4.60E+06	7.35E+07	6.78E+05
1	7.84E+07	1.89E+06	4.22E+07	1.55E+06	5.94E+07	2.15E+06
1.5	6.43E+07	6.05E+06	4.53E+07	2.89E+06	4.84E+07	7.11E+06
2	5.19E+07	4.08E+06	4.53E+07	2.04E+05	4.32E+07	2.90E+06
2.5	4.86E+07	4.51E+06	4.34E+07	2.19E+06	4.12E+07	4.39E+06
3	5.09E+07	6.44E+05	4.74E+07	6.57E+06	4.00E+07	1.63E+06
4	4.93E+07	2.66E+06	4.50E+07	1.88E+06	3.94E+07	1.33E+06
5	4.83E+07	3.26E+06	4.93E+07	4.31E+06	3.56E+07	1.80E+06
6	5.26E+07	1.10E+06	5.03E+07	5.22E+06	3.69E+07	8.52E+05

GdnHCl concentration (M)	$Em\ Max_{HSA\ 295\ nm}$	STDEV	$Em\ Max_{NATA\ 295\ nm}$	STDEV	$Em\ Max_{HSA\ 340\ nm}$	STDEV
0	327	1.15	348	0.71	422	2.52
0.5	329	1.15	346	0.71	422	0.71
1	330	1.15	347	0.71	421	1.53
1.5	332	0.00	347	0.00	416	0.00
2	341	1.00	346	2.12	419	2.08
2.5	342	2.08	348	2.12	421	2.52
3	344	1.00	348	0.00	416	1.15
4	345	0.58	349	0.00	416	2.65
5	343	1.00	346	2.83	421	2.52
6	343	1.41	346	1.41	421	1.53

Table 5.2: Monitoring HSA unfolding by GdnHCl using HSA Trp fluorescence and ProCharTS luminescence. 25 µM HSA was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. The area under the curve and Em. λ_{max} of HSA Trp fluorescence, NATA fluorescence, and ProCharTS luminescence in the presence of 0—6 M GdnHCl concentration were tabulated along with standard deviation.

HSA Trp fluorescence and ProCharTS luminescence values were used to obtain denaturation midpoints by nonlinear regression analysis. Trp fluorescence values (fluorescence intensity, emission maxima, and steady-state anisotropy) fit well to the Boltzmann function with R^2 values between 0.982—0.996 (Figure 5.4A, 5.4C, and 5.4E). The estimated denaturation midpoints obtained from the analysis of Trp fluorescence values were in the 1.20—1.71 M range. In the presence of 0.5 M GdnHCl, the Trp fluorescence intensity showed deviation from the fit line, possibly detecting the molten globule-like state. HSA ProCharTS luminescence intensity fit the exponential decay function with R^2 value of 0.985 (Figure 5.4B), but ProCharTS emission maxima were found to be insensitive to protein unfolding (Figure 5.4D). The denaturation midpoint estimated from ProCharTS luminescence was 1.07 M. Based on the denaturation midpoint values obtained, disruption of charged amino acid contacts leading to reduced charge recombination luminescence was predicted to happen earlier than Trp solvent exposure and global protein

unfolding. A marginal increase in denaturation midpoint from 1.2 to 1.3 M was observed when Trp fluorescence intensity at 0.5 M GdnHCl was excluded from nonlinear regression analysis.



`

Figure 5.4: Nonlinear regression analysis was used to estimate the denaturation midpoint of HSA unfolding by GdnHCl using Trp fluorescence and ProCharTS luminescence. HSA Trp fluorescence values fit the Boltzmann function and ProCharTS luminescence intensity data fit the exponential decay function. The fits for HSA Trp integrated fluorescence intensity (Figure 5.4A), integrated ProCharTS luminescence intensity (Figure 5.4B), Trp Em Max (Figure 5.4C), ProCharTS Em Max (Figure 5.4D), and Trp anisotropy (Figure 5.4E) were plotted. Denaturation

midpoint and R^2 values of obtained fits are reported in the figure. The absence of change in ProCharTS emission max (Figure 5.4D) with increasing GdnHCl concentration was shown by a horizontal dashed line.

Trp r_{ss} values were used to characterize the HSA unfolding pathway. The decrease in steady-state fluorescence anisotropy values upon protein unfolding is shown in Table 5.3. Trp r_{ss} decreased from 0.205 to 0.093 upon chemical denaturation in the presence of 6 M GdnHCl. NATA r_{ss} values were not sensitive to an increase in denaturant concentration.

GdnHCl concentration (M)	r_{ss} HSA 295 nm	STDEV	r_{ss} NATA 295 nm	STDEV
0	0.205	0.0187	0.036	0.0004
0.5	0.202	0.0172	0.041	0.0026
1	0.190	0.0115	0.040	0.0057
1.5	0.159	0.0144	0.039	0.0002
2	0.120	0.0024	0.039	0.0024
2.5	0.109	0.0098	0.040	0.0051
3	0.098	0.0120	0.035	0.0013
4	0.095	0.0078	0.041	0.0031
5	0.096	0.0095	0.039	0.0068
6	0.093	0.0094	0.039	0.0053

Table 5.3: Monitoring HSA unfolding by GdnHCl using Trp steady-state fluorescence anisotropy. 5 µM HSA and 10 µM NATA were incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. Steady-state fluorescence anisotropy data collected with Ex. λ = 295 nm was tabulated along with the observed standard deviation.

5.1.3 The application of dansyl fluorescence to monitor HSA-Dansyl unfolding

Dansyl steady-state fluorescence, steady-state anisotropy, and shift in emission maxima upon excitation at λ = 340 nm were used to monitor HSA-Dansyl unfolding. The primary amines of Lys sidechains in HSA were conjugated non-specifically with extrinsic chromophore dansyl. To avoid the possibility of homo-FRET because of the presence of multiple Lysines in HSA, the dye/protein labelling ratio (0.596) was kept low. A high steady-state fluorescence anisotropy value of 0.249 upon dansyl excitation at 340 nm was used to verify the success of HSA labelling with dansyl (Table 5.4).

Sample	$Abs_{280\ nm}$ corrected	Conc protein (µM)	Conc Dye (µM)	D/P ratio	r_{ss}
HSA-Dansyl	0.292	8.01	4.78	0.596	0.249

Table 5.4: HSA labelling with extrinsic chromophore dansyl. The primary amines in HSA Lysines were conjugated with extrinsic chromophore dansyl with an efficiency of 59.6 %. The absorbance of 20 times diluted labelled protein was used to calculate concentration and D/P ratio, and measure steady-state anisotropy.

HSA-Dansyl unfolding caused a decrease in dansyl fluorescence intensity (Figure 5.5A and 5.5B), a decrease in dansyl fluorescence anisotropy (Figure 5.5D), and a red shift in emission maxima (Figure 5.5C), with the data obtained suggesting complete protein unfolding at 3.5 M GdnHCl

concentration. The decrease in fluorescence intensity and red shift were attributed to dansyl solvent exposure upon protein unfolding. The increase of dansyl fluorescence intensity in the presence of 1 M GdnHCl hinted to the presence of previously reported molten globule-like state of HSA in sub-denaturing GdnHCl concentrations. HSA-Dansyl unfolding was performed in the presence of a reducing agent (0.5 mM TCEP) to reduce the disulfide bonds and ensure complete protein unfolding.

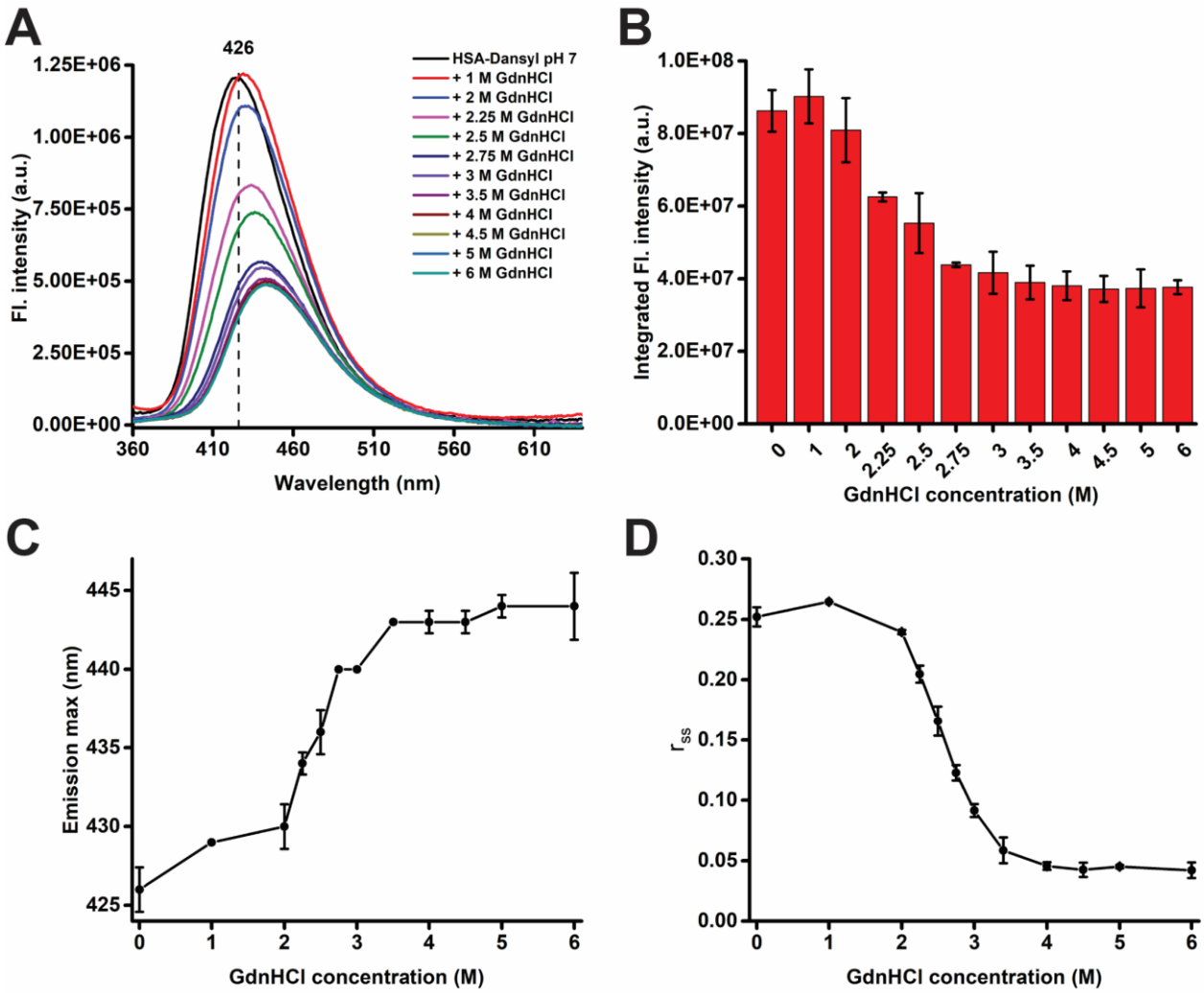


Figure 5.5: Monitoring HSA-Dansyl unfolding by GdnHCl using dansyl steady-state fluorescence. 10 µM HSA-Dansyl was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. Steady-state dansyl fluorescence was collected in the range of 360—620 nm (Figure 5.5A) with Ex. λ = 340 nm using Ex. slit = 2 nm and Em. slit = 5 nm. Dansyl area under the fluorescence curve (Figure 5.5B) and Em. λ_{max} (Figure 5.5C) were also reported. Steady-state fluorescence anisotropy was collected with Ex. λ = 340 nm and Em. λ = 425 nm using Ex. slit = 2 nm and Em. slit = 10 nm. Dansyl steady-state anisotropy with increasing denaturant concentration was represented as a scatter plot (Figure 5.5D).

HSA-Dansyl denaturation midpoints were obtained by performing nonlinear regression analysis of dansyl fluorescence data collected upon HSA-Dansyl unfolding. The data fit well to the

Boltzmann function with R^2 values between 0.983—0.999 (Figure 5.6A—C). The denaturation midpoints obtained by dansyl fluorescence were between 2.24—2.58 M. Slight deviation of dansyl integrated fluorescence intensity from the fitted curve was observed at 1 M GdnHCl concentration. Upon exclusion of dansyl fluorescence intensity at 1 M GdnHCl from nonlinear curve fitting, estimated denaturation midpoint increased marginally from 2.24 to 2.32 M.

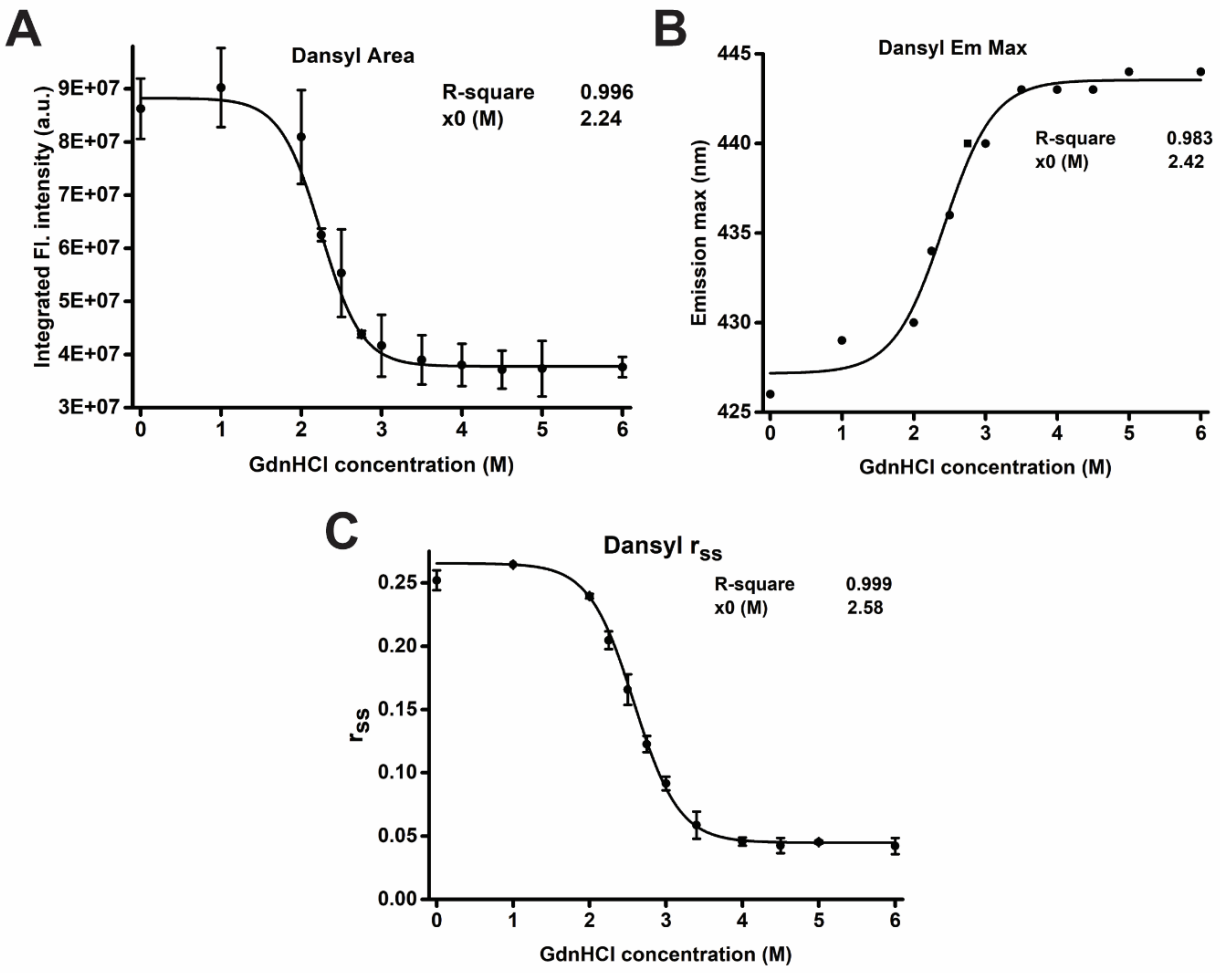


Figure 5.6: Nonlinear regression analysis was used to estimate the denaturation midpoint of HSA-Dansyl unfolding by GdnHCl using dansyl fluorescence and anisotropy. HSA-Dansyl integrated fluorescence intensity (Figure 5.6A), Em Max (Figure 5.6B), and steady-state anisotropy (Figure 5.6C) values fit the Boltzmann function. Denaturation midpoint and R^2 values of obtained fits are reported in the figure.

The dansyl steady-state fluorescence, emission maxima, and anisotropy values were monitored upon HSA-Dansyl unfolding by GdnHCl (Table 5.5). The dansyl area under the curve decreased by 56.4 %, and the emission maximum shifted from 426 to 444 nm in the presence of 6 M GdnHCl. The dansyl anisotropy reached a baseline value at 4 M GdnHCl. The increased dansyl area under the curve at 1 M GdnHCl concentration allowed the detection of molten globule-like intermediate with extrinsic chromophore dansyl.

GdnHCl concentration (M)	$Area_{Dansyl}$	STDEV	Em Max_{Dansyl}	STDEV	$r_{ss\ Dansyl}$	STDEV
0	8.63E+07	5.70E+06	426	1.41	0.252	0.0079
1	9.02E+07	7.43E+06	429	0.00	0.264	0.0001
2	8.09E+07	8.81E+06	430	1.41	0.240	0.0018
2.25	6.25E+07	1.20E+06	434	0.71	0.205	0.0070
2.5	5.53E+07	8.24E+06	436	1.41	0.166	0.0121
2.75	4.38E+07	6.28E+05	441	0.71	0.123	0.0064
3	4.17E+07	5.82E+06	440	0.00	0.091	0.0054
3.5	3.90E+07	4.63E+06	443	0.00	0.059	0.0107
4	3.80E+07	3.98E+06	443	0.71	0.046	0.0032
4.5	3.72E+07	3.58E+06	443	0.71	0.042	0.0060
5	3.73E+07	5.23E+06	444	0.71	0.045	0.0013
6	3.76E+07	1.92E+06	444	2.12	0.042	0.0065

Table 5.5: Monitoring HSA-Dansyl unfolding by GdnHCl using dansyl steady-state fluorescence. 10 µM HSA-Dansyl was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. The area under the fluorescence curve, Em. λ_{max}, and steady-state anisotropies were tabulated along with the observed standard deviation.

Nonlinear regression analysis was used to estimate the denaturation midpoints of HSA unfolding. The estimated denaturation midpoints varied significantly with the technique used, and differences in chromophore were reflected in the midpoint values obtained. The midpoint values were estimated to be from 0.61—2.58 M (Figure 5.7). Although there was considerable variation between chromophores, it was observed that data for a particular chromophore fit the same function, and the estimated denaturant midpoint values were also in good agreement with each other. HSA absorbance values fit the exponential decay function with the midpoint values estimated between 0.61—0.80 M. ProCharTS luminescence intensity data also fit the exponential decay function whereas MRE data fit the logistic function. Trp and dansyl fluorescence data fit the Boltzmann function. HSA denaturation midpoint analysis suggested that charged amino acid contacts got disrupted earlier than Trp exposure to solvent and loss of secondary structure elements upon protein unfolding. Of the chromophores used to monitor HSA unfolding, dansyl exposure to solvent was the slowest and happened at the highest denaturant concentration. Trp fluorescence anisotropy and MRE values are most sensitive to the global unfolding status of protein, and the estimated denaturation midpoints were thus in good agreement with each other.

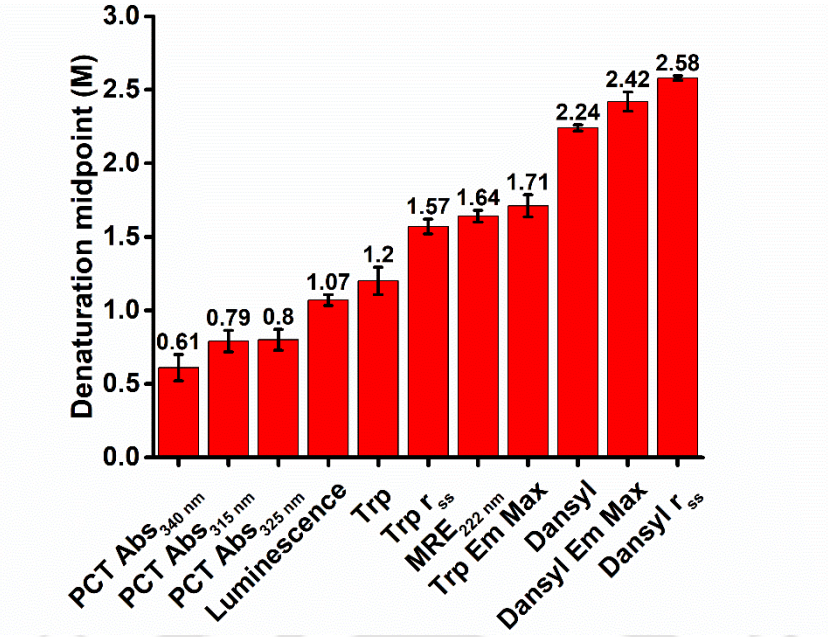


Figure 5.7: Estimated denaturation midpoint values of HSA unfolding using ProCharTS and conventional methods. Curve fitting was done with inbuilt functions in Origin software.

The quantitative analysis of the HSA unfolding pathway was done by multiple techniques to obtain the denaturation midpoints (Table 5.6). The estimated denaturation midpoints from HSA ProCharTS absorbance and luminescence were low, suggesting that disruption of charged amino acid contacts happened early in the protein unfolding process.

	Denaturation midpoint (M)	Std Error	R^2
$Abs_{340\ nm}$	0.61	0.090	0.999
$Abs_{315\ nm}$	0.79	0.073	0.983
$Abs_{325\ nm}$	0.80	0.071	0.991
Luminescence	1.07	0.037	0.985
Trp	1.20	0.093	0.984
Trp r_{ss}	1.57	0.050	0.996
$MRE_{222\ nm}$	1.64	0.041	0.999
Trp Em Max	1.71	0.076	0.982
Dansyl fluo Area	2.24	0.021	0.996
Dansyl Em Max	2.42	0.065	0.983
Dansyl fluo r_{ss}	2.58	0.016	0.999

Table 5.6: Parameters obtained after GdnHCl induced HSA and HSA-Dansyl denaturation midpoint analysis. The denaturation midpoint, the standard error in predicting the midpoint, and R^2 values were tabulated.

5.1.4 The application of protein charge transfer absorbance spectra to monitor urea induced HSA unfolding

Decrease in HSA ProCharTS absorbance at multiple wavelengths was observed after urea induced unfolding (**Figure 5.8A**). A pronounced decrease in HSA absorbance was noticeable from 6 to 8

M samples compared to lower urea concentrations (**Figure 5.8B**). The decrease in HSA ProCharTS absorbance upon urea induced unfolding additionally confirmed that the absorbance decrease in GdnHCl experiments was not due to neutralization of charges by the guanidine/chloride ions used to induce protein unfolding.

Urea significantly affects Trp absorbance and the presence of urea caused a red shift of HSA Trp absorption spectrum in the near UV region (inset of **Figure 5.8A**). The observed red shift affected the ProCharTS in the 295—325 nm range, causing an abrupt absorbance increase in 1.5 M urea (which correlated with a red shift of Trp absorbance in 1.5 M and higher urea concentrations). HSA Trp absorbance maximum was red-shifted by ~11 nm (from 278 to 289 nm) in 8 M urea. Quantitative analysis of HSA unfolding by urea was performed at 340, 370, and 400 nm to avoid contribution from red-shifted Trp absorbance values (**Figure 5.9A—C**). Regardless of the effects of urea on Trp absorbance, the decrease in ProCharTS absorbance upon protein unfolding at higher denaturant concentrations could be established. HSA unfolding was also evident in the CD spectra collected at increasing denaturant concentration and the increase in $MRE_{222\ nm}$ (**Figure 5.8C**) is consistent with HSA unfolding.

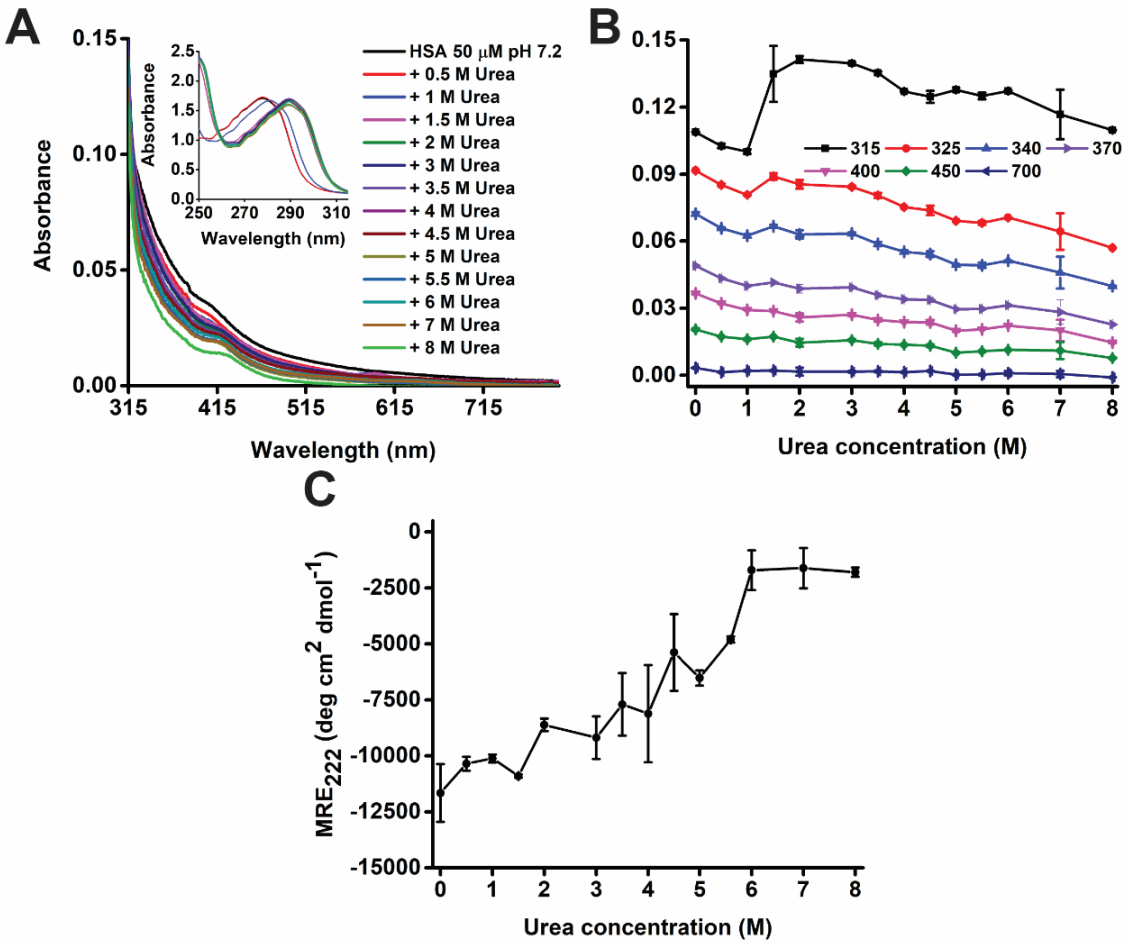


Figure 5.8: Monitoring HSA unfolding by urea using ProCharTS and CD spectroscopy. 50 µM HSA was incubated in 0—8 M urea (20 mM NaH_2PO_4, 0.5 mM TCEP, pH 7.2) for 12 hours at 25 °C. Absorbance spectra was recorded between 250—800 nm (**Figure 5.8A**) for three

independent samples and the absorbance at multiple wavelengths was monitored upon HSA unfolding and represented as a scatter plot (**Figure 5.8B**). The concurrent ProCharTS and aromatic amino acid absorbance in the 250—315 nm region was plotted in the inset of **Figure 5.8A**. CD spectra was recorded for samples diluted to 0.5 µM HSA concentration and the $MRE_{222\ nm}$ was reported for two independent measurements (**Figure 5.8C**).

When HSA unfolding was induced by urea, ProCharTS absorbance values fit the exponential decay function (**Figure 5.9A—C**) whereas MRE values fit the logistic function (**Figure 5.9D**). Denaturation midpoints obtained upon ProCharTS absorbance analysis (1.58—2.64 M) were lower than $MRE_{222\ nm}$ (5.48 M) analysis. This indicated that disruption of charged amino acid contacts preceded the loss of secondary structure.

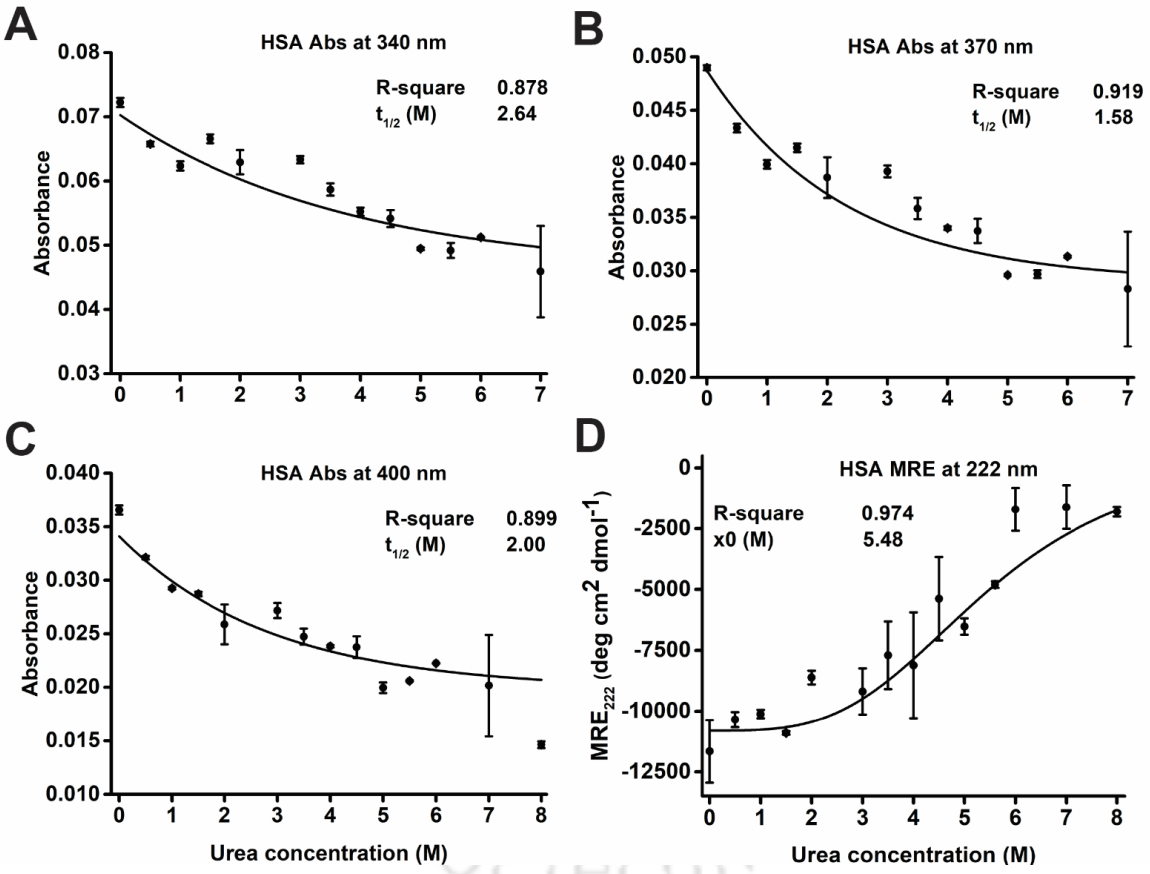


Figure 5.9: Nonlinear regression analysis was used to estimate the denaturation midpoint of HSA unfolding by urea using HSA ProCharTS absorbance and MRE. ProCharTS absorbance at 340 nm (**Figure 5.9A**), 370 nm (**Figure 5.9B**), 400 nm (**Figure 5.9C**) fit the exponential decay function while the $MRE_{222\ nm}$ fit the logistic function (**Figure 5.9D**). Denaturation midpoint and R^2 values of obtained fits are reported in the figure. Curve fitting for ProCharTS at 340 and 370 nm was successful only after excluding the absorbance value at 8 M urea.

5.1.5 The application of protein charge transfer absorbance spectra to monitor α_3W unfolding

The chemical denaturation of α_3W by GdnHCl was performed, and protein unfolding was monitored using α_3W ProCharTS absorbance at multiple wavelengths (Figure 5.10A) and far UV CD spectroscopy (Figure 5.10C). 25 µM α_3W unfolding caused a substantial decrease in ProCharTS, with $Abs_{340\ nm}$ decreasing from 0.030 to 0.009 in the presence of 6 M GdnHCl (Figure 5.10B). This decrease in α_3W ProCharTS was observed at multiple wavelengths upon protein unfolding. α_3W ProCharTS decrease was detected even in the presence of 0.5 M GdnHCl, with the absorbance decreasing to a baseline value in the 3 M GdnHCl sample. The disruption of charged amino acid sidechain contacts due to protein unfolding resulted in the decreased ProCharTS. α_3W $MRE_{222\ nm}$ increase upon chemical denaturation was additionally used to monitor protein unfolding. The mean residual ellipticity at 222 nm started to increase in the presence of 1.5 M GdnHCl, and complete unfolding was observed at 4 M GdnHCl (Figure 5.10D). In the absence of denaturant, α_3W was rich in helical content, which decreased with protein unfolding. Upon complete unfolding, α_3W became predominantly unstructured. Nonlinear regression analysis was used to estimate the denaturation midpoint of α_3W unfolding.

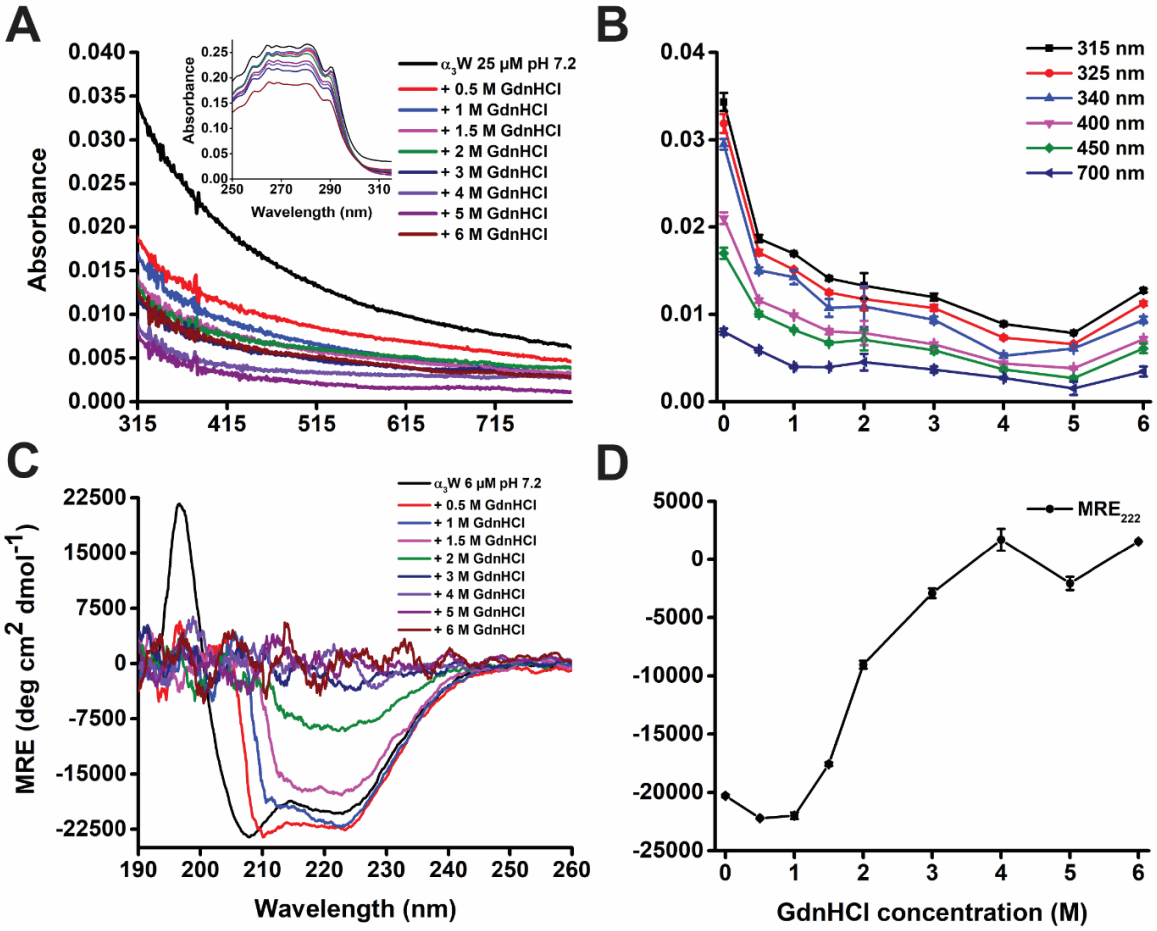


Figure 5.10: Monitoring α_3W unfolding by GdnHCl using ProCharTS and CD spectroscopy. 25 µM α_3W was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, pH 7.2) for 12 hours at 25 °C. Absorbance spectra were recorded between 250—800 nm (Figure 5.10A) for three independent

samples, and the absorbance at multiple wavelengths was monitored upon α_3W unfolding and represented as a scatter plot (Figure 5.10B). CD spectra were recorded for samples diluted to 6 µM α_3W concentration in the 190—260 nm region (Figure 5.10C), and the $MRE_{222\ nm}$ was reported for two independent measurements (Figure 5.10D). The concurrent ProCharTS and aromatic amino acid absorbance in the 250—315 nm region was plotted in the inset of Figure 5.10A.

The simultaneous decrease in α_3W ProCharTS at 315, 325, and 340 nm with an increase in $MRE_{222\ nm}$ values was used to infer the disruption of charged amino acid sidechain contacts and loss of secondary structure due to protein unfolding (Table 5.7). α_3W CD spectra were collected in the presence of increasing GdnHCl concentration, and the analysis of secondary structure parameters was performed. α_3W helical content decreased substantially in the presence of 1.5 M GdnHCl, with complete unfolding being observed at 4 M GdnHCl. Upon complete unfolding, α_3W was predominantly a random coil. The quantitative analysis of α_3W unfolding pathway was performed by nonlinear regression analysis of the absorbance and MRE values.

GdnHCl concentration (M)	$Abs_{315\ nm}$	STDEV	$Abs_{325\ nm}$	STDEV	$Abs_{340\ nm}$	STDEV	$MRE_{222\ nm}$	STDEV	α-helix	β-sheet	Random coil
0	0.034	0.0010	0.032	0.0011	0.030	0.0006	-20290	112	94.78	0.06	5.16
0.5	0.019	0.0004	0.017	0.0003	0.015	0.0004	-22209	46	87.90	0.03	12.07
1	0.017	0.0001	0.015	0.0001	0.014	0.0008	-21997	276	77.50	0.06	22.44
1.5	0.014	0.0002	0.013	0.0002	0.011	0.0010	-17575	144	65.64	0.40	34.32
2	0.013	0.0015	0.012	0.0013	0.011	0.0023	-9039	338	28.91	12.01	59.08
3	0.012	0.0004	0.011	0.0004	0.010	0.0000	-2913	419	2.05	32.76	65.19
4	0.009	0.0002	0.007	0.0002	0.005	0.0003	1697	945	0.81	38.50	60.69
5	0.008	0.0001	0.007	0.0001	0.006	0.0003	-2064	571	1.25	34.03	64.72
6	0.013	0.0002	0.011	0.0003	0.009	0.0004	1556	5	0.90	37.62	61.48

Table 5.7: Monitoring α_3W unfolding by GdnHCl using ProCharTS and CD spectroscopy. 25 µM α_3W was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, pH 7.2) for 12 hours at 25 °C. $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ with increasing denaturant concentration upon α_3W unfolding were collected and tabulated. CD spectra were recorded for samples diluted to 6 µM concentration, and the $MRE_{222\ nm}$ values were reported. Secondary structure analysis parameters obtained by K2D3 software upon α_3W unfolding were also tabulated.

Information about the denaturation midpoints was obtained by the nonlinear regression analysis of absorbance and MRE values at multiple points in the α_3W unfolding pathway. $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ values fit well to the exponential decay function with R^2 values between 0.900—0.908 (Figure 5.11 A—C). $MRE_{222\ nm}$ fit the logistic function with a high R^2 value of 0.999 (Figure 5.11D). Based on the denaturation midpoints obtained upon monitoring α_3W unfolding using ProCharTS and $MRE_{222\ nm}$, the disruption of charged amino acid contacts was predicted to occur earlier than the loss of secondary structure. The midpoint obtained from $Abs_{325\ nm}$ and $Abs_{340\ nm}$ analysis were 0.86 and 0.66 M, compared to 1.99 M for $MRE_{222\ nm}$ analysis.

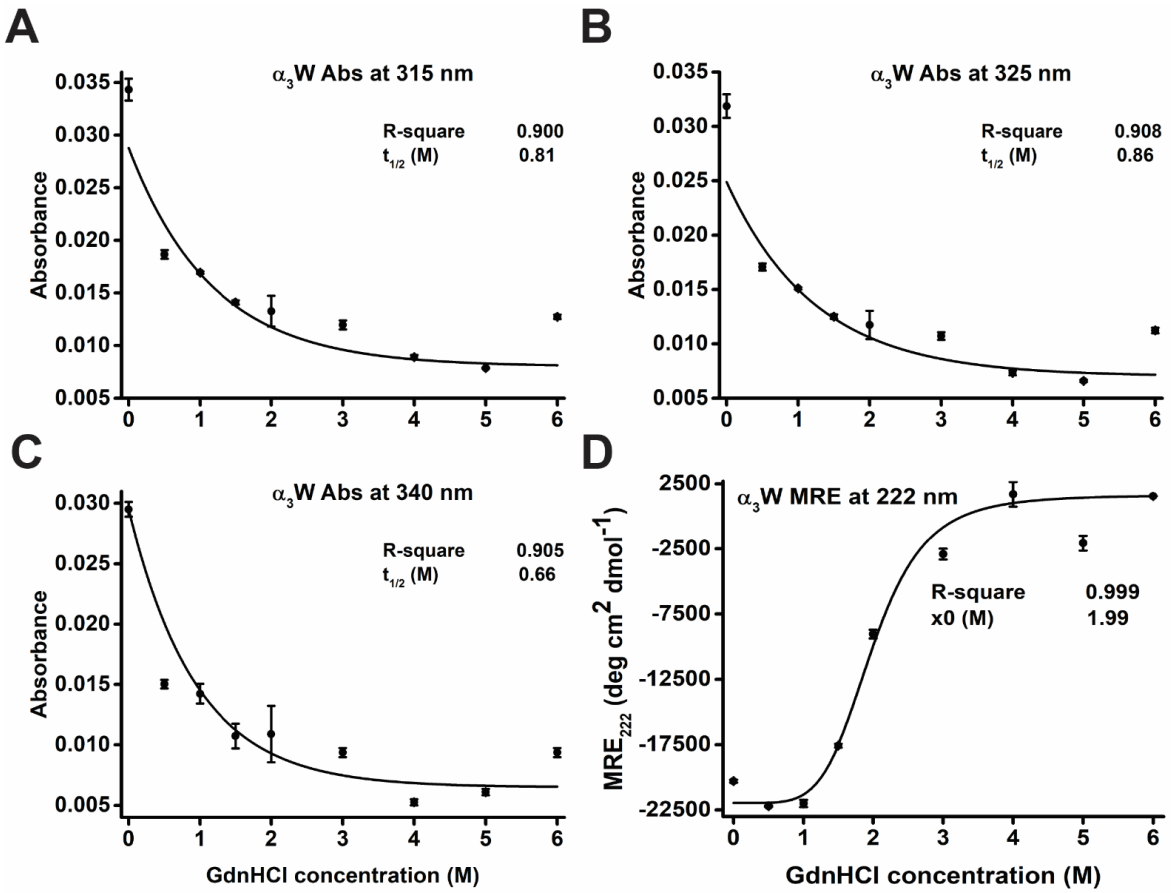


Figure 5.11: Nonlinear regression analysis was used to estimate the denaturation midpoint of α3W unfolding by GdnHCl using ProCharTS and CD spectroscopy. α_3W absorbance values at 315 nm (Figure 5.11A), 325 nm (Figure 5.11B), and 340 nm (Figure 5.11C) fit the exponential decay function, whereas MRE data (Figure 5.11D) fit the logistic function. Denaturation midpoint and R^2 values of obtained fits are reported in the figure.

5.1.6 The application of Trp fluorescence spectra to monitor α3W unfolding

α_3W Trp emission at λ_{ex} = 295 nm was used to monitor α_3W unfolding by GdnHCl. α_3W unfolding caused a decrease in Trp fluorescence (Figure 5.12A and 5.12B), a decrease in Trp r_{ss} (Figure 5.12D), and a red shift in Trp emission maxima (Figure 5.12C). α_3W Trp fluorescence rose at low GdnHCl concentration but decreased as expected in the presence of 2 M GdnHCl and reached a baseline value at 3 M GdnHCl. The red shift in Trp emission maxima and decrease in Trp r_{ss} at 1.5 M GdnHCl correlated well with the increase observed in $MRE_{222\ nm}$. Trp emission maximum reached a baseline value at 3 M GdnHCl, while Trp r_{ss} was the lowest at 4 M GdnHCl concentration. Due to low quantum yield, α_3W ProCharTS luminescence could not be used to monitor protein unfolding. The denaturation midpoint for α_3W unfolding was estimated using nonlinear regression analysis (Figure 5.13). NATA fluorescence intensity, emission maxima, and steady-state fluorescence anisotropy did not change in the presence of chemical denaturant.

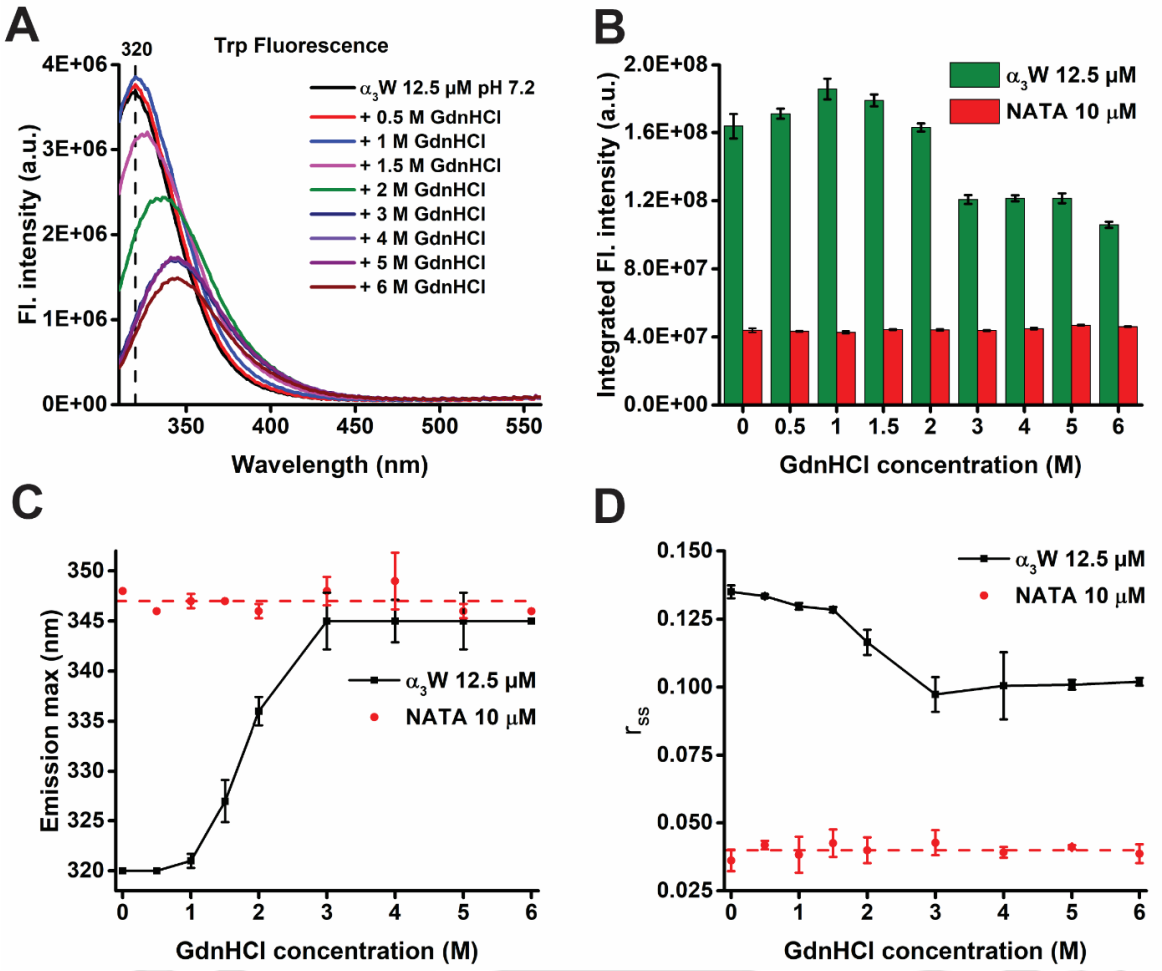


Figure 5.12: Monitoring α3W unfolding by GdnHCl using Trp fluorescence and anisotropy. 12.5 μM α3W was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, pH 7.2) for 12 hours at 25 °C. Steady-state fluorescence was collected for 12.5 μM α3W and 10 μM NATA with Ex. λ = 295 nm and emission in the range of 310—560 nm (Figure 5.12A) using Ex. slit = 2 nm and Em. slit = 5 nm. Area under the curve (Figure 5.12B) and Em. λ_{max} (Figure 5.12C) were plotted. Steady-state anisotropy was collected (Figure 5.12D) with Ex. λ = 295 nm and Em. λ = 320 nm using Ex. slit = 2 nm and Em. slit = 15 nm. The horizontal dashed lines represent the absence of change in NATA emission max (Figure 5.12C) and NATA anisotropy (Figure 5.12D) with increasing GdnHCl concentration.

Trp fluorescence intensity decrease and red shift in emission maxima were used to monitor α3W unfolding. Trp fluorescence intensity, emission maxima, and anisotropy values at various points in the α3W unfolding pathway are listed in Table 5.8. Trp fluorescence intensity decreased by 26.2 % in the presence of 3 M GdnHCl. Trp emission maximum shifted from 320 to 345 nm, suggesting Trp solvent exposure upon protein unfolding. Decrease in Trp r_{ss} was also observed upon α3W unfolding. NATA fluorescence intensity values were insensitive to increasing denaturant concentration as expected.

GdnHCl concentration (M)	$Area_{\alpha3W\ 295\ nm}$	STDEV	$Area_{NATA\ 295\ nm}$	STDEV	$Em\ Max_{\alpha3W\ 295\ nm}$	STDEV	$Em\ Max_{NATA\ 295\ nm}$	STDEV
0	1.64E+08	7.30E+06	4.39E+07	1.22E+06	320	0.00	348	0.00
0.5	1.71E+08	3.08E+06	4.33E+07	3.61E+05	320	0.00	346	0.00
1	1.86E+08	6.13E+06	4.27E+07	7.74E+05	321	0.71	347	0.71
1.5	1.79E+08	3.49E+06	4.43E+07	4.84E+05	327	2.12	347	0.00
2	1.63E+08	2.40E+06	4.42E+07	5.96E+05	336	1.41	346	0.71
3	1.21E+08	2.60E+06	4.38E+07	4.92E+05	345	2.83	348	1.41
4	1.21E+08	1.78E+06	4.49E+07	6.12E+05	345	2.12	349	2.83
5	1.21E+08	2.90E+06	4.70E+07	3.57E+05	345	2.83	346	0.71
6	1.06E+08	1.80E+06	4.60E+07	2.82E+05	345	0.00	346	0.00

GdnHCl concentration (M)	$r_{ss\ \alpha3W\ 295\ nm}$	STDEV	$r_{ss\ NATA\ 295\ nm}$	STDEV
0	0.135	0.0024	0.036	0.0040
0.5	0.133	0.0006	0.042	0.0016
1	0.130	0.0012	0.038	0.0066
1.5	0.128	0.0010	0.043	0.0051
2	0.116	0.0046	0.040	0.0047
3	0.097	0.0064	0.043	0.0046
4	0.101	0.0123	0.039	0.0020
5	0.101	0.0018	0.041	0.0007
6	0.102	0.0014	0.039	0.0035

Table 5.8: Monitoring α_3W unfolding by GdnHCl using Trp fluorescence and anisotropy. 12.5 µM α_3W was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, pH 7.2) for 12 hours at 25 °C. Steady-state fluorescence was collected for 12.5 µM α_3W and 10 µM NATA with Ex. λ = 295 nm in the range of 310—560 nm using Ex. slit = 2 nm and Em. slit = 5 nm. The Em. λ_{max} and area under the curve were tabulated. Steady-state anisotropy was collected with Ex. λ = 295 nm and Em. λ = 320 nm using Ex. slit = 2 nm and Em. slit = 15 nm and tabulated along with the observed standard deviation.

Nonlinear regression analysis was used to obtain α_3W unfolding denaturation midpoints using Trp fluorescence and anisotropy. α_3W Trp integrated fluorescence intensity (Figure 5.13A), α_3W Trp emission maxima (Figure 5.13B), and anisotropy (Figure 5.13C) fit the Boltzmann function with R^2 values of 0.940, 0.999 and 0.992 respectively. Denaturation midpoints of 2.40, 1.81, and 2.05 M were obtained by analysing integrated fluorescence intensity, emission maxima and steady-state anisotropy values. Trp solvent exposure was predicted to occur earlier than global protein unfolding (measured by Trp anisotropy) based on denaturation midpoint analysis. Due to low quantum yield, α_3W ProCharTS luminescence was not employed to monitor protein unfolding and subsequent denaturation midpoint analysis.

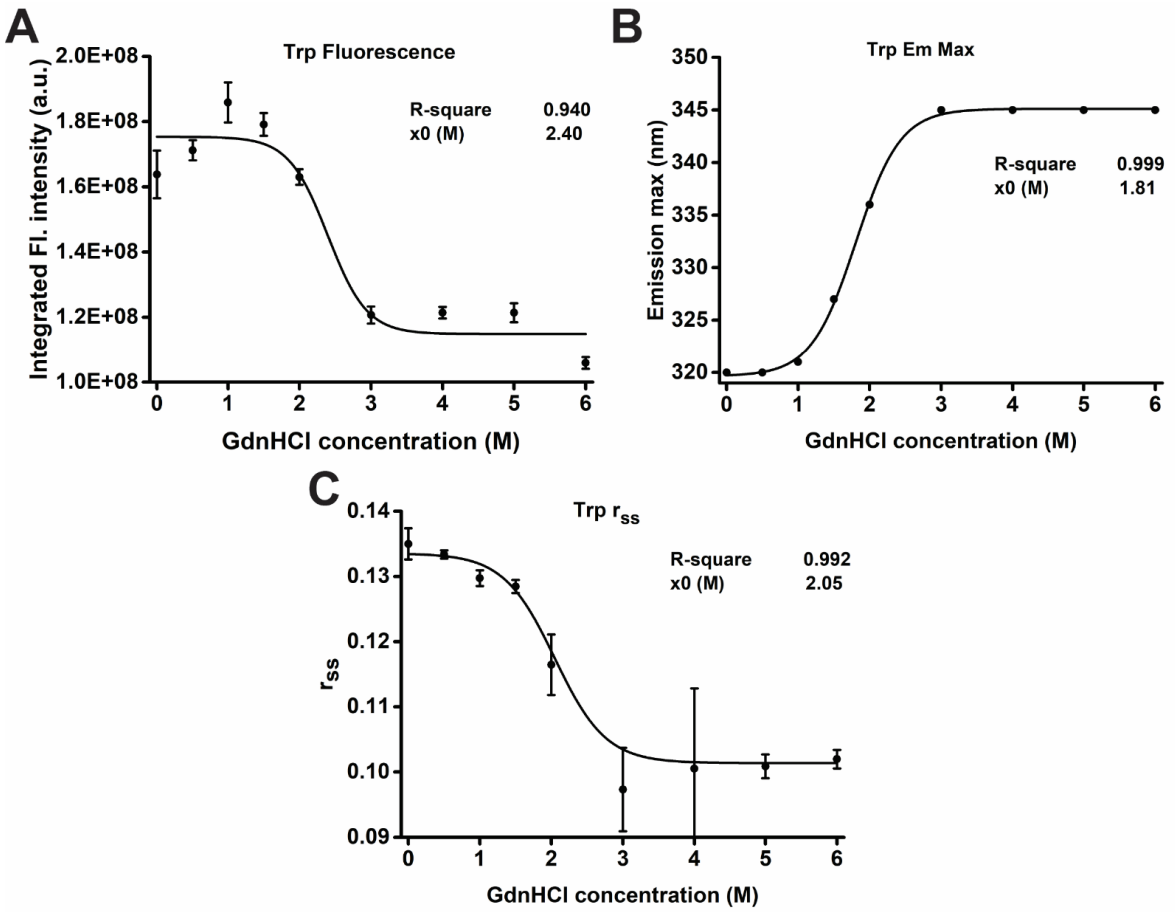


Figure 5.13: Nonlinear regression analysis was used to estimate the denaturation midpoint of α_3W unfolding by GdnHCl using Trp fluorescence and anisotropy. α_3W Trp integrated fluorescence intensity (Figure 5.13A), Trp emission maxima (Figure 5.13B), and anisotropy (Figure 5.13C) fit the Boltzmann function. Denaturation midpoint and R^2 values of obtained fits are reported in the figure.

α_3W unfolding was performed, and nonlinear regression analysis was used to estimate the denaturation midpoints using multiple techniques (Figure 5.14). Trp r_{ss} and $MRE_{222\ nm}$ are most sensitive to global protein unfolding, and the obtained denaturation midpoints were similar. The midpoint values were estimated to be between 0.66—0.86 M. α_3W absorbance values fit the exponential decay function whereas MRE values fit the Logistic function. Trp fluorescence values fit the Boltzmann function. α_3W denaturation midpoint analysis suggested that loss of charged amino acid contacts preceded Trp solvent exposure. Loss of secondary structure elements and increase in rotational diffusion occurred at higher denaturant concentration.

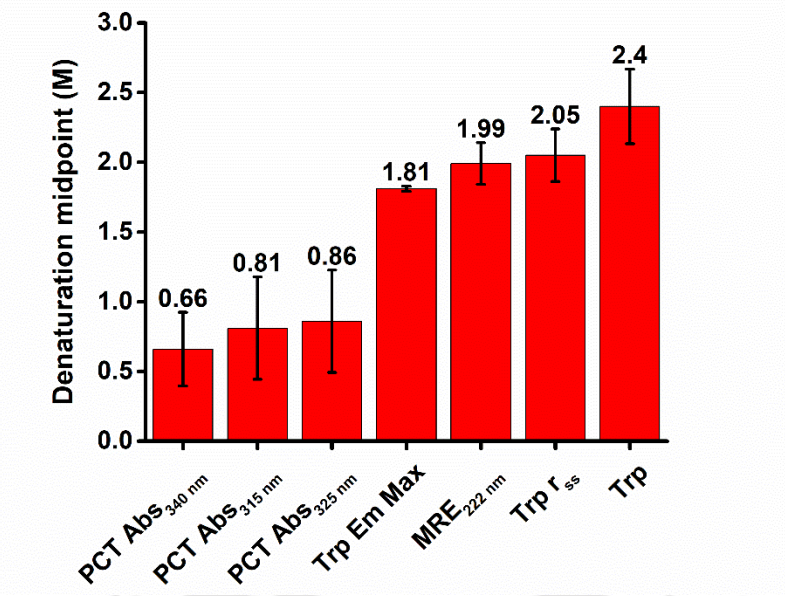


Figure 5.14: Parameters obtained for GdnHCl induced α_3W denaturation midpoint analysis. Nonlinear regression analysis was done to determine the denaturation midpoint by employing the absorbance, Trp steady-state fluorescence, and anisotropy at multiple points in the unfolding pathway. Curve fitting was done with inbuilt functions in Origin software.

α_3W unfolding pathway was characterized quantitatively by multiple techniques to obtain the denaturation midpoints (Table 5.9). The estimated denaturation midpoints obtained upon analysis of α_3W ProCharTS absorbance pointed to the disruption of charged amino acid contacts early in the protein unfolding process.

	Denaturation midpoint (M)	Std Error	R^2
$Abs_{340\ nm}$	0.66	0.264	0.905
$Abs_{315\ nm}$	0.81	0.367	0.900
$Abs_{325\ nm}$	0.86	0.367	0.908
Trp Em Max	1.81	0.018	0.999
$MRE_{222\ nm}$	1.99	0.150	0.999
Trp r_{ss}	2.05	0.189	0.992
Trp Area	2.42	0.306	0.939

Table 5.9: Parameters obtained after GdnHCl induced α_3W denaturation midpoint analysis. The denaturation midpoint, the standard error in predicting the midpoint, and R^2 values were tabulated.

5.1.7 The application of protein charge transfer absorbance spectra to monitor α_3C unfolding

α_3C protein unfolding was performed with chemical denaturant GdnHCl and monitored using ProCharTS and CD spectroscopy (Figure 5.15A and 5.15C). α_3C ProCharTS decrease was employed to follow GdnHCl-induced protein unfolding. α_3C $Abs_{325\ nm}$ decreased from 0.018 to 0.005 upon complete unfolding by 6 M GdnHCl (Figure 5.15B). α_3C ProCharTS decreased substantially in the presence of 0.5 M GdnHCl, and the absorbance decreased to a baseline value in 3 M GdnHCl. $MRE_{222\ nm}$ was used to track α_3C unfolding, which increased significantly at and

beyond 2 M GdnHCl concentration (Figure 5.15D). Due to low quantum yield, α_3C luminescence could not be used to track protein unfolding. Nonlinear regression analysis was used to estimate α_3C unfolding denaturation midpoints (Figure 5.16).

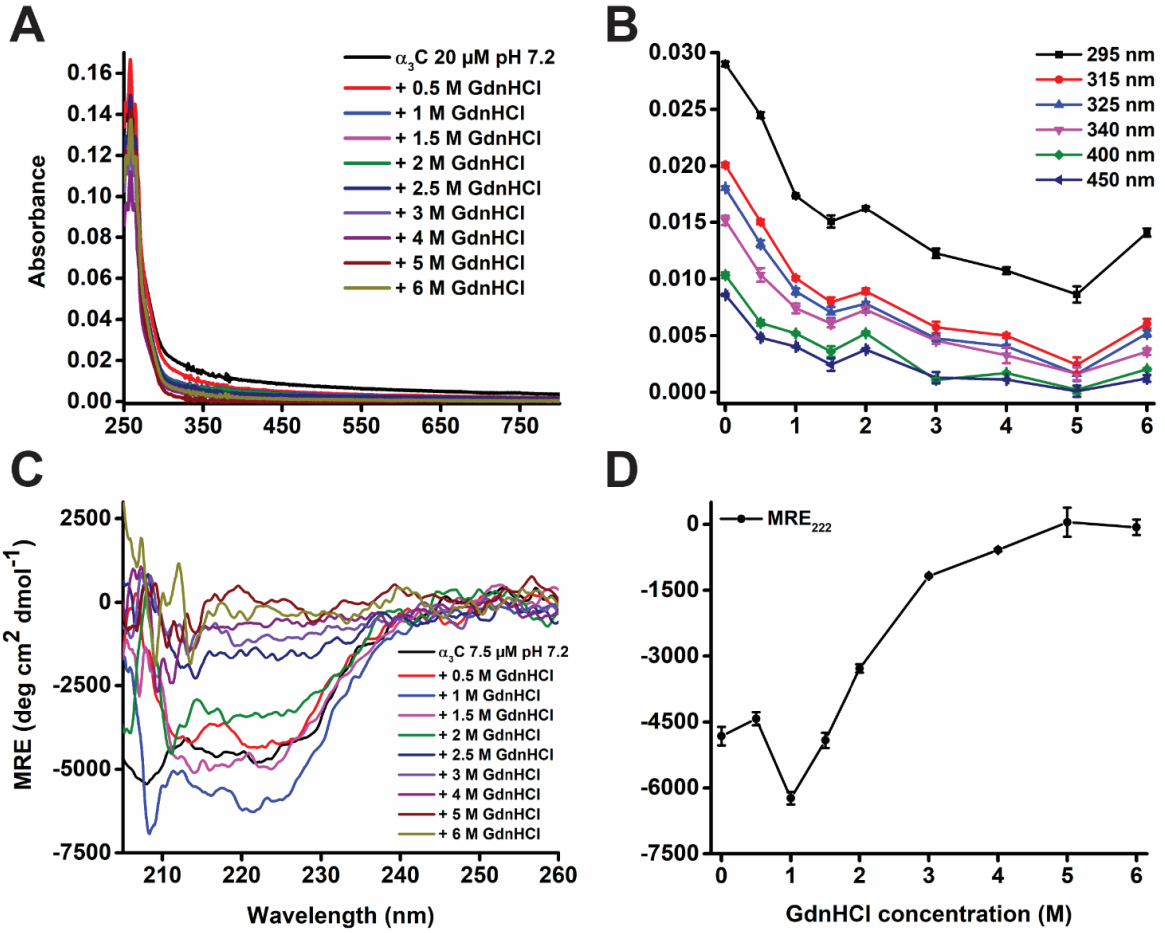


Figure 5.15: Monitoring α_3C unfolding by GdnHCl using ProCharTS and CD spectroscopy. 20 µM α_3C was incubated in 0—6 M GdnHCl (20 mM Tris, pH 7.2) for 12 hours at 25 °C. Absorbance spectra were recorded in the 250—800 nm region (Figure 5.15A). The absorbance at multiple wavelengths were monitored with increasing denaturant concentration during unfolding and plotted (Figure 5.15B). CD spectra were recorded for samples diluted to 7.5 µM concentration in the 200—260 nm region (Figure 5.15C), and the $MRE_{222\ nm}$ reported (Figure 5.15D).

α_3C ProCharTS absorbance and MRE values were collected along various points in the GdnHCl-induced protein unfolding pathway. Nonlinear regression analysis was used to estimate the denaturation midpoints for α_3C unfolding. $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ values fit well to the exponential decay function with R^2 values of 0.973, 0.976, and 0.916 respectively (Figure 5.16A—C), and $MRE_{222\ nm}$ fit the logistic function with a R^2 value of 0.965 (Figure 5.16D). The midpoints obtained from $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ analysis were 0.71, 0.88, and 0.96 M, which was much lower than the value of 2.16 M obtained from $MRE_{222\ nm}$ analysis. Thus the disruption of α_3C charged amino acid sidechain contacts happened much earlier than the loss of secondary structure.

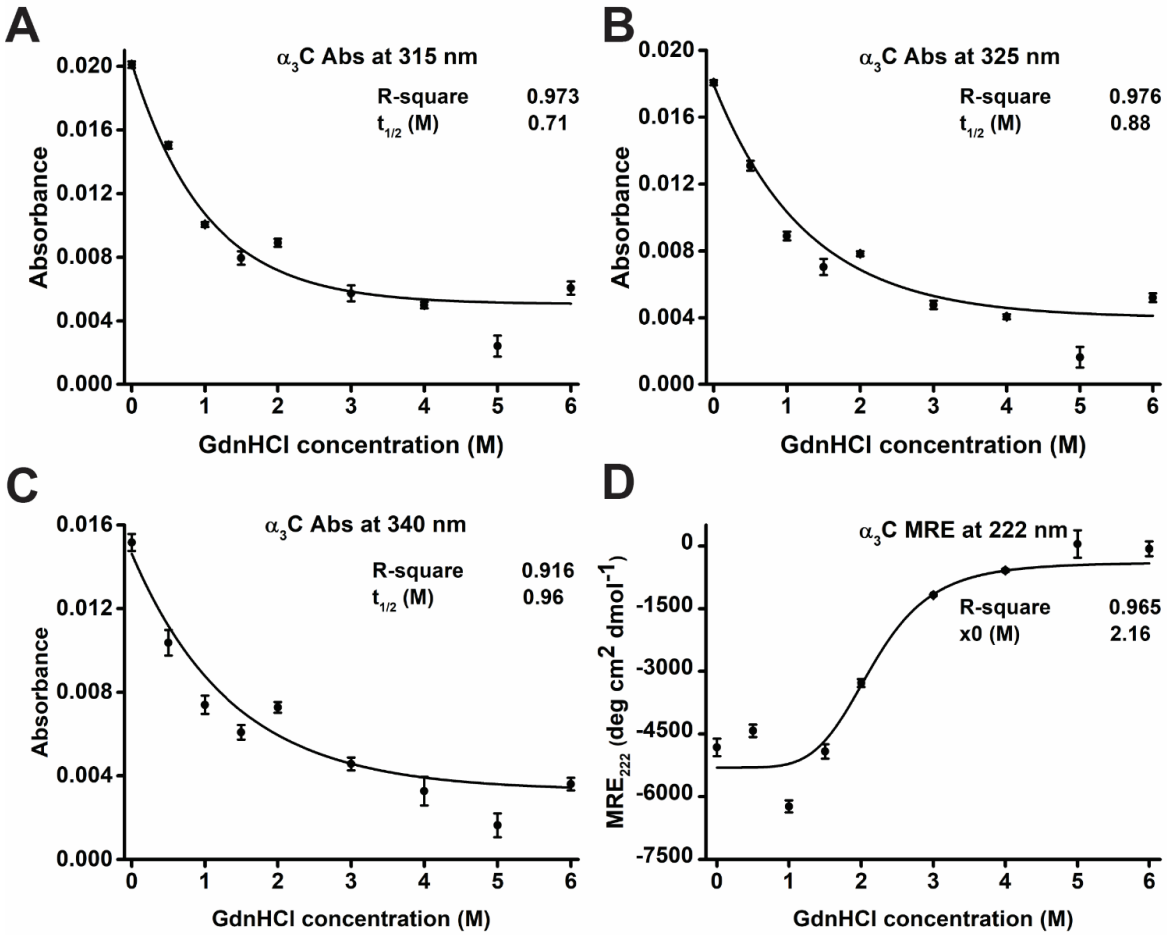


Figure 5.16: Nonlinear regression analysis was used to estimate the denaturation midpoint of α3C unfolding by GdnHCl using ProCharTS and CD spectroscopy. α3C ProCharTS at 315 nm (Figure 5.16A), 325 nm (Figure 5.16B), and 340 nm (Figure 5.16C) fit the exponential decay function. MRE values (Figure 5.16D) fit the logistic function. Denaturation midpoint and R^2 values of obtained fits are reported in the figure.

The decrease in α3C ProCharTS values at 315, 325, 340 nm and simultaneous increase in $MRE_{222\ nm}$ values were used to monitor charged amino acid sidechain contact disruption and secondary structure loss upon protein unfolding (Table 5.10). Secondary structure analysis was performed on α3C CD spectra collected at increasing GdnHCl concentration. Helix to coil transition was observed upon α3C unfolding. α3C helicity decreased in the presence of 1.5 M GdnHCl, with complete unfolding being observed at 3 M GdnHCl concentration. Due to significant absorbance by Tris buffer in the 190—200 nm region, detector saturation was observed in the collected CD spectra. The secondary structure analysis was affected by this, and the helix content obtained was lesser than upon similar analysis for α3W. However, the parameters obtained were used to conclude the helix to coil transition upon protein unfolding. The actual α3C helix content is expected to be much higher (92.32 % helix content was observed for folded α3C in 20 mM NaH_2PO_4, pH 7 buffer). Nonlinear regression analysis of the absorbance and MRE values collected during the α3C unfolding process was used to obtain denaturation midpoints (Figure 5.17).

GdnHCl concentration (M)	$Abs_{315\ nm}$	STDEV	$Abs_{325\ nm}$	STDEV	$Abs_{340\ nm}$	STDEV	$MRE_{222\ nm}$	STDEV	α-helix	β-sheet	Random coil
0	0.020	0.0002	0.018	0.0002	0.015	0.0004	-4822	211	43.74	7.25	49.01
0.5	0.015	0.0002	0.013	0.0003	0.010	0.0006	-4425	149	41.7	17.1	41.2
1	0.010	0.0002	0.009	0.0003	0.007	0.0004	-6234	144	49	10.02	40.98
1.5	0.008	0.0004	0.007	0.0005	0.006	0.0003	-4918	172	28.7	19.51	51.79
2	0.009	0.0003	0.008	0.0001	0.007	0.0003	-3281	96	18.6	19.43	61.97
3	0.006	0.0005	0.005	0.0003	0.005	0.0003	-1175	19	0.73	44.56	54.71
4	0.005	0.0002	0.004	0.0002	0.003	0.0007	-585	31	0.99	40.99	58.02
6	0.006	0.0004	0.005	0.0003	0.004	0.0003	-68	176	0.49	40.86	58.65

Table 5.10: Monitoring α_3C unfolding by GdnHCl using ProCharTS and CD spectroscopy. 20 µM α_3C was incubated in 0—6 M GdnHCl (20 mM Tris, pH 7.2) for 12 hours at 25 °C. $Abs_{315\ nm}$, $Abs_{325\ nm}$, and $Abs_{340\ nm}$ at increasing denaturant concentration during α_3C unfolding were tabulated. CD spectra were recorded for 7.5 µM α_3C, and the $MRE_{222\ nm}$ was reported. Secondary structure analysis parameters obtained by K2D3 software upon α_3C unfolding were tabulated.

The denaturation midpoint estimated by analysing ProCharTS absorbance during α_3C unfolding was between 0.71—0.96 M compared to 2.16 M from MRE analysis (Figure 5.17). α_3C ProCharTS at 315 nm, 325 nm, and 340 nm fit the exponential decay model whereas MRE fit the logistic model. The significant difference in predicted denaturation midpoint by ProCharTS and far UV CD spectroscopy pointed towards the earlier disruption of charged amino contacts than the loss of secondary structure. The observed difference illustrated the importance of comparing quantitative analysis of protein unfolding pathway by ProCharTS with conventional methods.

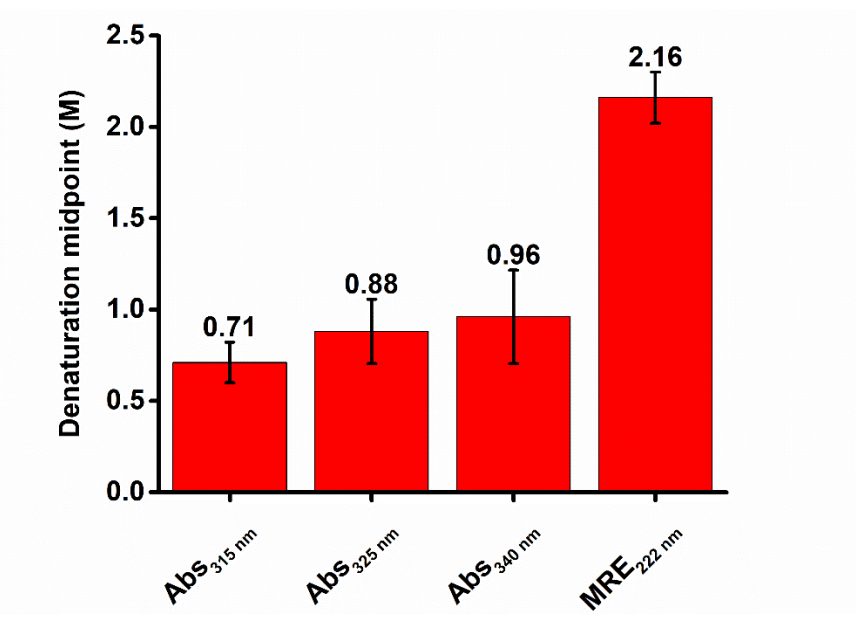


Figure 5.17: Parameters obtained for GdnHCl induced α_3C denaturation midpoint analysis. Nonlinear regression analysis was done to determine the denaturation midpoints by employing the ProCharTS and MRE values at multiple points in the unfolding pathway. Curve fitting was done with inbuilt functions in the Origin software.

Denaturation midpoints of α_3C unfolding pathway were estimated by nonlinear regression analysis of α_3C ProCharTS and MRE values and reported in Table 5.11. The midpoint values suggested that disruption of charged amino acid contacts preceded the loss of secondary structure. Thus ProCharTS absorbance was employed in the quantitative analysis of α_3C protein unfolding.

	Denaturation midpoint (M)	Std Error	R^2
$Abs_{315\ nm}$	0.71	0.111	0.973
$Abs_{325\ nm}$	0.88	0.176	0.976
$Abs_{340\ nm}$	0.96	0.256	0.916
$MRE_{222\ nm}$	2.16	0.14	0.965

Table 5.11: Parameters obtained after GdnHCl induced α_3C denaturation midpoint analysis. The denaturation midpoint, the standard error in predicting the midpoint, and R^2 values were tabulated.

5.1.8 The application of dansyl fluorescence to monitor α_3C-Dansyl unfolding

α_3C-Dansyl unfolding was attempted to be monitored using changes in dansyl steady-state fluorescence, steady-state anisotropy, and a possible shift in emission maxima upon protein unfolding. Site-specific labelling of α_3C single Cys sidechain was done with extrinsic chromophore dansyl. The efficiency of labelling reaction was 86.0 %, and the higher steady-state fluorescence anisotropy value of 0.015 (Table 5.12) confirmed α_3C labelling with dansyl (r_{ss} of free dye fraction was 0.004). The r_{ss} value was lower compared to dansyl in HSA-Dansyl (Table 5.4) due to the smaller size of α_3C.

Sample	$Abs_{280\ nm}$ corrected	Conc protein (μM)	Conc Dye (μM)	D/P ratio	r_{ss}
α_3C-Dansyl	0.0063	2.07	1.78	0.860	0.015

Table 5.12: α_3C labelling with extrinsic chromophore Dansyl. α_3C single Cys sidechain was conjugated with extrinsic chromophore dansyl with an efficiency of 86.0 %. The absorbance of 4 times diluted labelled protein was used to calculate concentration, D/P ratio, and measure steady-state anisotropy.

Absence of change in dansyl fluorescence and anisotropy was observed upon α_3C-Dansyl unfolding. An increase in chemical denaturant concentration and resultant protein unfolding did not change the dansyl fluorescence intensity (Figure 5.18A and 5.18B), position of emission maxima (Figure 5.18C), and r_{ss} (Figure 5.18D). The insensitivity of dansyl fluorescence to protein unfolding was attributed to the chromophore in α_3C-Dansyl possibly being solvent-exposed even in the folded protein. This was illustrated by a red-shifted emission maximum of 504 nm upon excitation at 340 nm compared to 426 nm for HSA-Dansyl. The decrease in dansyl fluorescence, red shift of emission maxima, and decrease in anisotropy observed for HSA-Dansyl (Figure 5.5) upon protein unfolding was not detected upon α_3C-Dansyl unfolding. Thus ProCharTS absorbance proved to be more useful in monitoring α_3C unfolding compared to the extrinsic chromophore dansyl.

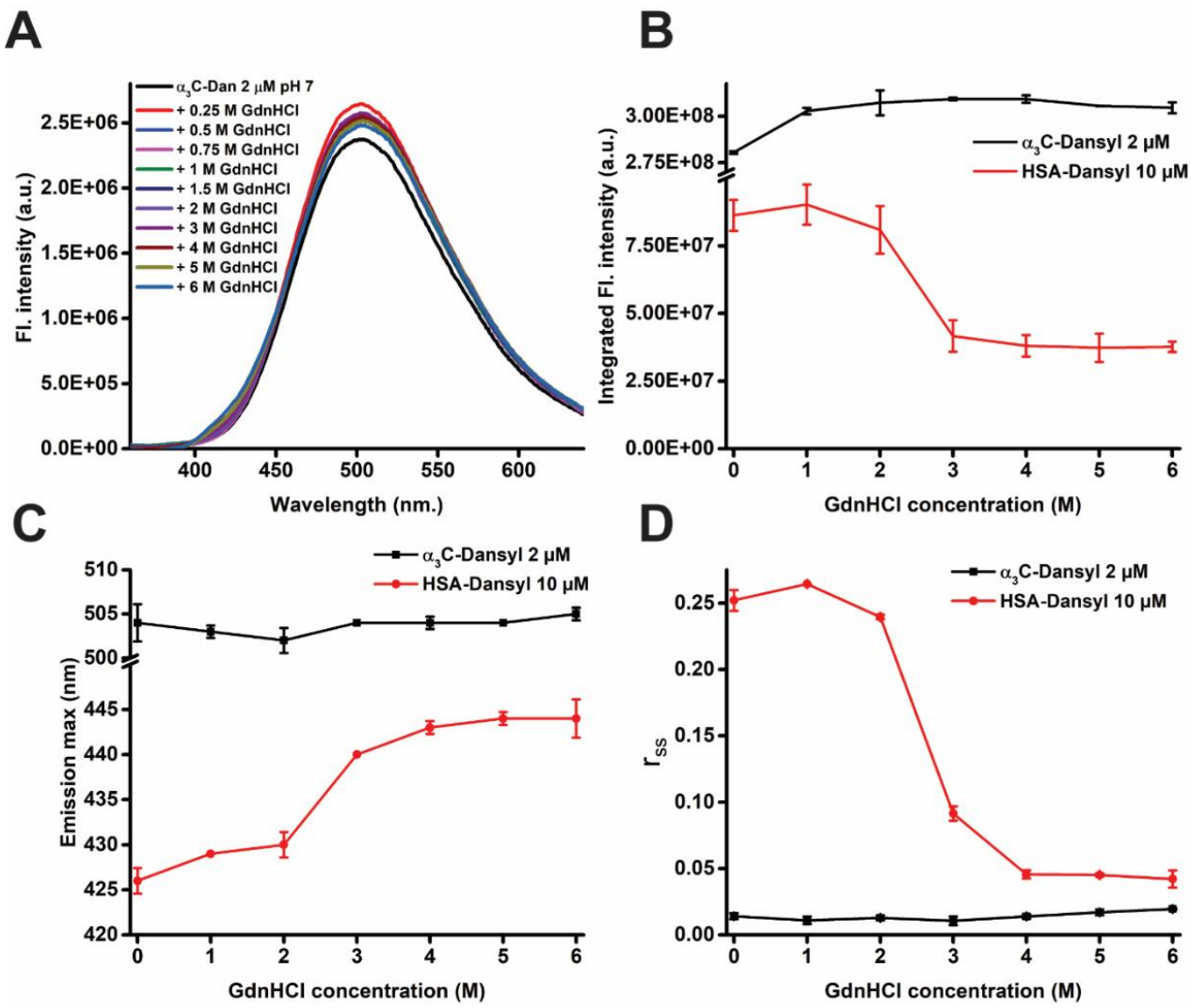


Figure 5.18: Monitoring α3C-Dansyl unfolding by GdnHCl using dansyl steady-state fluorescence. 2 μM α_3C-Dansyl was incubated in 0—6 M GdnHCl (20 mM NaH_2PO_4, pH 7.2) for 12 hours at 25 °C. Steady-state fluorescence was collected with Ex. λ = 340 nm and emission in the range of 360—620 nm using Ex. slit = 2 nm and Em. slit = 15 nm (Figure 5.18A). The area under the curve (Figure 5.18B) and Em. λ_{max} (Figure 5.18C) was also monitored. Steady-state anisotropy was collected with Ex. λ = 340 nm and Em. λ = 500 nm using Ex. slit = 2 nm and Em. slit = 15 nm (Figure 5.18D). α_3C-Dansyl fluorescence was compared to HSA-Dansyl fluorescence (data corresponding to Figure 5.5).

5.2 Conclusions

1. HSA unfolding pathway was characterized using the novel intrinsic ProCharTS absorbance and luminescence. HSA unfolding was also monitored using the conventional methods of CD spectroscopy, Trp steady-state fluorescence, and Dansyl steady-state fluorescence.
2. α_3W unfolding was characterized using ProCharTS absorbance, Trp fluorescence, Trp anisotropy, and CD spectroscopy.
3. α_3C unfolding was monitored using a decrease in ProCharTS absorbance and an increase in mean residual ellipticity at 222 nm. ProCharTS was more useful in tracking α_3C unfolding compared to extrinsic chromophore dansyl.
4. Nonlinear regression analysis was used to determine HSA, α_3C, and α_3W unfolding denaturation midpoints. Based on the estimated midpoints, the disruption of charged amino acid sidechain contacts was predicted to happen earlier than loss of secondary structure elements and Trp solvent exposure during the protein unfolding process.
5. ProCharTS absorbance was thus applied for the LABEL-FREE detection of HSA, α_3W, and α_3C unfolding.

CHAPTER 6

MONITORING DNA-PROTEIN BINDING USING PROTEIN CHARGE TRANSFER SPECTRA

6.1 Monitoring DNA-Protein binding using protein charge transfer spectra

The physiological role of PRM as a DNA condensing agent is well known. The non-specific binding of PRM to multiple types of DNA and subsequent DNA condensation is also known. In the present study, UV-Vis absorbance, fluorescence spectroscopy, circular dichroism spectroscopy and electrophoretic mobility shift assays were used to monitor DNA-protein binding. The binding of HEWL and PRM with calf thymus genomic DNA (gDNA) was monitored using gDNA absorbance, HEWL absorbance, HEWL Trp fluorescence, and PRM ProCharTS absorbance. DNA-protein binding was additionally confirmed by monitoring precipitation of nucleoprotein complex in electrophoretic mobility shift assays. The absence of DNA condensation by BSA was also observed as reported earlier.[20]

6.1.1 Monitoring the binding of HEWL with gDNA using DNA absorbance

The ability of HEWL to bind DNA *in-vitro* has been reported earlier.[94,95] HEWL binding with DNA leads to the formation and subsequent precipitation of nucleoprotein complex. In the present study, the precipitation of HEWL-gDNA nucleoprotein complex was used to monitor DNA-HEWL binding spectrophotometrically.

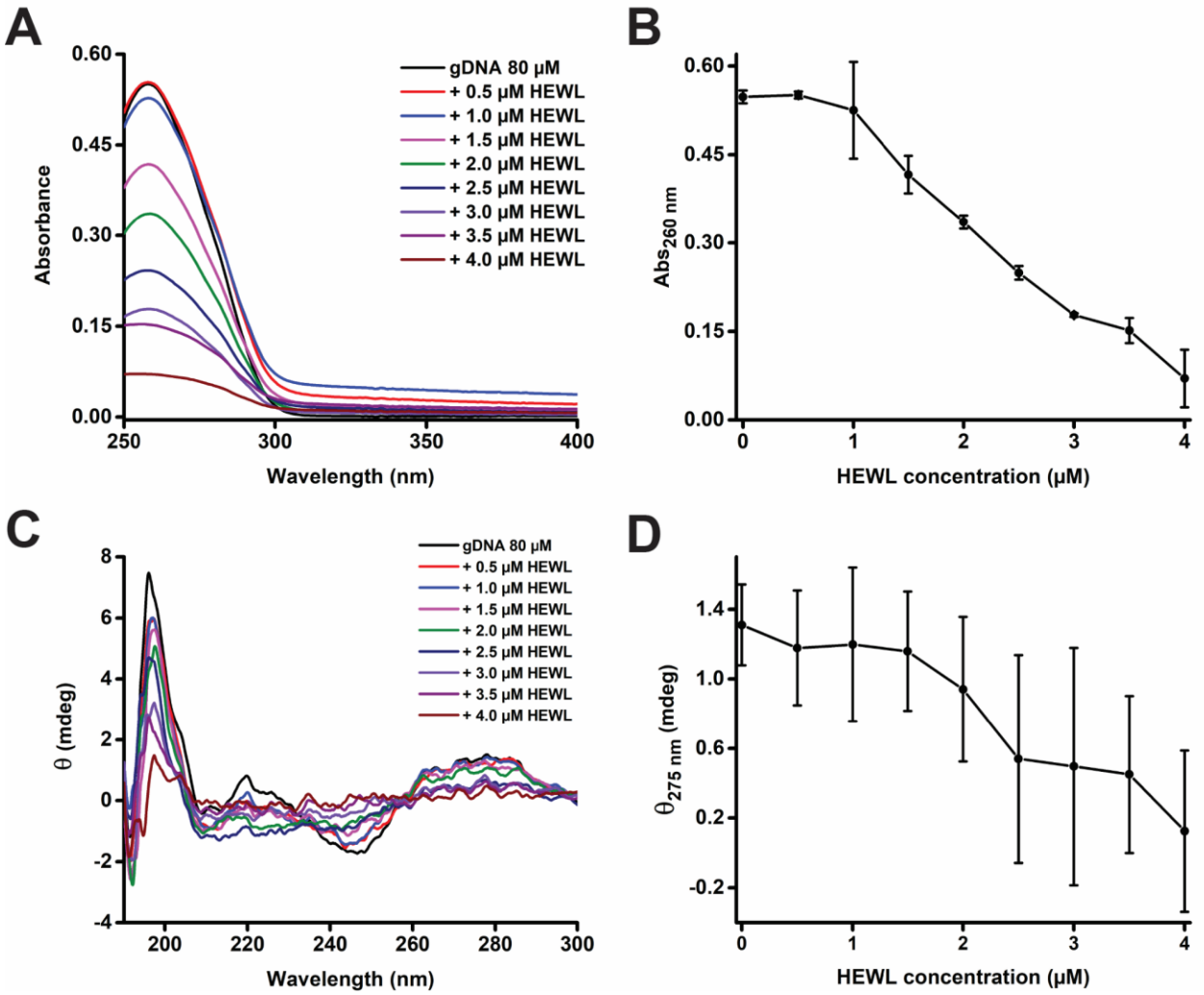
The DNA condensation assay was performed and nucleoprotein complex precipitation was monitored using decrease in gDNA absorbance (Figure 6.1A). gDNA (effective nucleotide concentration of 80 µM) absorbance decreased even in the gDNA-HEWL mixture supernatant containing moderate amount of HEWL (1 µM). Nucleoprotein complex precipitation was observed to a greater extent at higher HEWL concentration. Higher concentrations of HEWL reduced the absorbance even further which was reflected in the gDNA absorbance decrease at 260 nm (Figure 6.1B). The decrease in DNA concentration of HEWL-gDNA mixture supernatants happened due to precipitation of the nucleoprotein complex. Almost complete DNA precipitation was observed in the HEWL-gDNA mixture containing 4 µM HEWL.

The decrease in DNA circular dichroism due to nucleoprotein complex precipitation was used to monitor HEWL-gDNA binding. Upon HEWL binding, DNA concentration in the DNA-protein mixture supernatants decreased due to precipitation of the nucleoprotein complex. The decreased supernatant DNA concentration resulted in reduced circular dichroism of the samples (Figure 6.1C). Due to weak chirality of B-form DNA, very low circular dichroism was observed but decrease in ellipticity at 275 nm with increasing HEWL concentration was apparent (Figure 6.1D).

The scattering of light by nucleoprotein complexes was used to monitor gDNA binding with HEWL. An increase in scattering of HEWL-gDNA mixtures (prior to centrifugation) was observed with increasing HEWL concentration (Figure 6.1E). The scattering monitored at 350 nm increased in the presence of 0.5—3.5 µM HEWL (Figure 6.1F) due to the presence of nucleoprotein precipitates in HEWL-gDNA mixtures. The increased scattering at higher HEWL concentrations implied the greater extent of nucleoprotein complex formation and subsequent precipitation at higher protein concentrations. In the presence of 4 µM protein, there was complete precipitation

of nucleoprotein complex which appeared as a white clump that settled to the bottom of the cuvette resulting in lower scattering.

The precipitation of gDNA by HEWL was additionally monitored using agarose gel electrophoresis (Figure 6.1G). DNA decrease in the supernatant of gDNA-HEWL mixtures was observed with increasing HEWL concentration. The intensity of the sheared DNA band (higher electrophoretic mobility band) decreased even in the presence of 0.5 µM HEWL and complete precipitation in gDNA + 4 µM HEWL sample was observed. The intensity of the low electrophoretic mobility band (corresponding to unsheared gDNA) remained similar in the 0—1.5 µM HEWL range. In the presence of low HEWL concentration, complete precipitation of the nucleoprotein complex is not expected. $Abs_{260\ nm}$ did not decrease in the presence of 0.5 µM HEWL, hinting that the nucleoprotein complex does not undergo complete precipitation at low protein concentration. At higher HEWL concentrations, precipitation of the low electrophoretic mobility band was also observed.



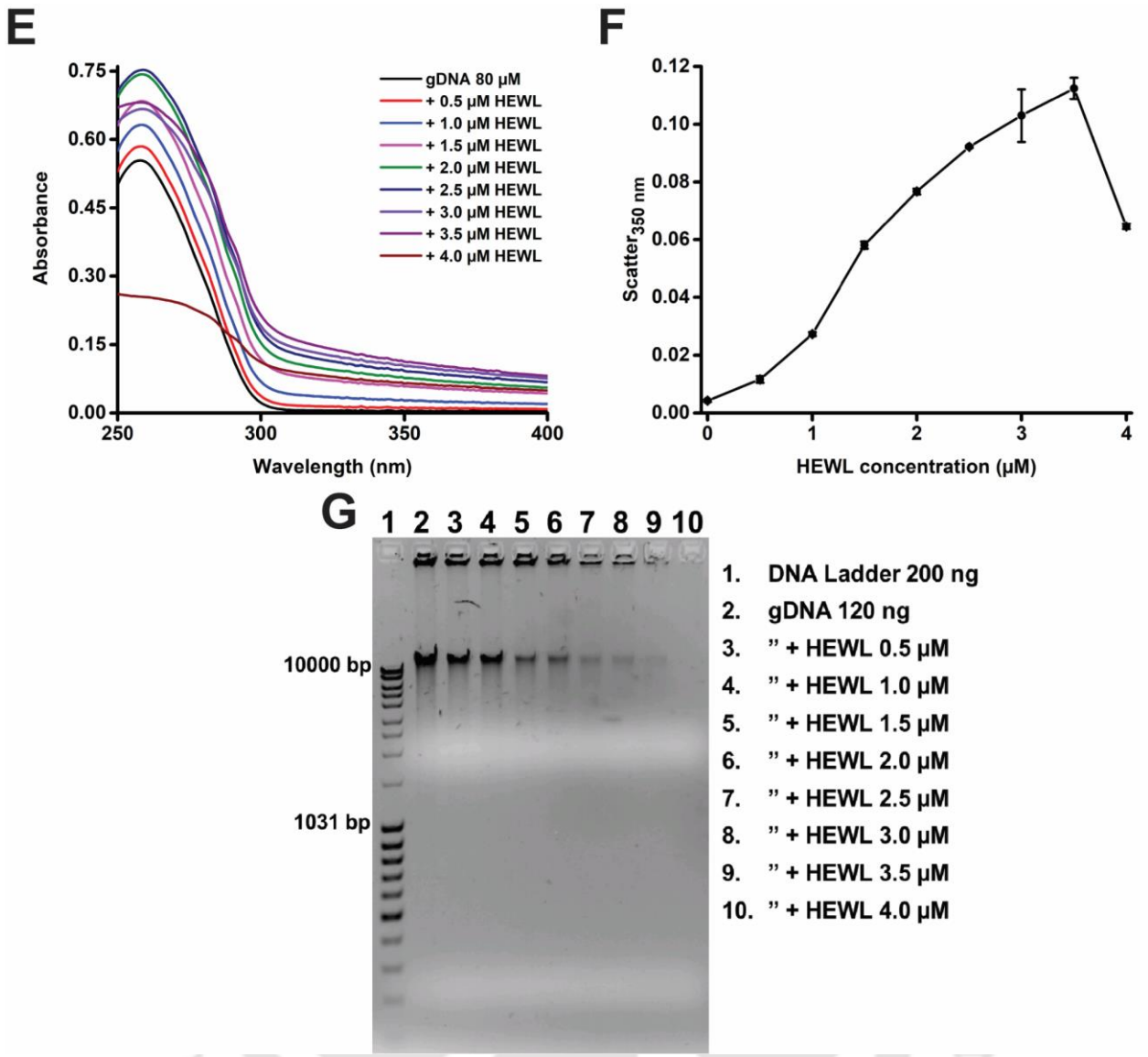


Figure 6.1: Monitoring the binding of HEWL with gDNA using DNA absorbance. The binding of gDNA and HEWL was monitored using DNA absorbance, scattering, and CD spectroscopy with increasing HEWL concentration. To a fixed concentration of gDNA (80 µM effective nucleotide concentration), increasing concentrations of HEWL (0.5—4 µM) were added. Sample preparation and assay were performed in 50 mM NaCl. DNA and HEWL mixing was done at 4 °C for 15 minutes at 120 rpm. After gDNA-HEWL mixing, centrifugation at 12000 rpm was done for 15 minutes. Absorbance was monitored between 250—400 nm (Figure 6.1A) and absorbance at 260 nm (Figure 6.1B) was plotted. CD spectra was measured between 190—300 nm (Figure 6.1C) and ellipticity at 275 nm (Figure 6.1D) of the supernatant was used to characterize DNA-HEWL binding. The scattering of nucleoprotein complex was collected for HEWL-gDNA mixtures before centrifugation (Figure 6.1E) and plotted at 350 nm (Figure 6.1F). The precipitation of nucleoprotein complex was visualized using 1 % agarose gel electrophoresis of the HEWL-gDNA mixture supernatants (Figure 6.1 G).

HEWL-gDNA binding was characterized by monitoring the increased scattering by nucleoprotein complex at 350 nm, decrease in DNA absorbance at 260 nm and decrease in circular dichroism at

275 nm (Table 6.1). $Abs_{260\ nm}$ decreased from 0.548 to 0.070 in the supernatant of HEWL-gDNA mixture with 4 µM HEWL. The ellipticity at 275 nm was very low and decreased from 1.310 mdeg to 0.125 mdeg. Almost no scattering at 350 nm was observed from 80 µM gDNA sample. HEWL-gDNA mixing led to nucleoprotein complex formation and precipitation with a concomitant increase in scattering from 0.004 to 0.112 in the presence of 3.5 µM HEWL. The scattering observed in the presence of 4 µM HEWL was an outlier.

HEWL concentration (µM)	$Abs_{260\ nm}$	STDEV	$\theta_{275\ nm}$	STDEV	$Scatter_{350\ nm}$	STDEV
0	0.548	0.0109	1.310	0.2317	0.004	0.0004
0.5	0.551	0.0059	1.178	0.3300	0.012	0.0012
1	0.525	0.0820	1.198	0.4415	0.027	0.0007
1.5	0.416	0.0320	1.159	0.3431	0.058	0.0013
2	0.336	0.0113	0.941	0.4160	0.077	0.0010
2.5	0.249	0.0115	0.540	0.5970	0.092	0.0004
3	0.178	0.0032	0.497	0.6821	0.103	0.0091
3.5	0.152	0.0213	0.451	0.4510	0.112	0.0037
4	0.070	0.0488	0.125	0.4630	0.065	0.0009

Table 6.1: Monitoring the binding of HEWL with gDNA using DNA absorbance. The binding of gDNA and HEWL was monitored using DNA absorbance at 260 nm, scattering at 350 nm, and DNA ellipticity at 275 nm (reported in mdeg) with increasing HEWL concentration and the data tabulated along with the standard deviation.

6.1.2 Monitoring the binding of gDNA with HEWL using HEWL absorbance and fluorescence

HEWL absorbance, Trp fluorescence, far UV CD spectroscopy, and agarose gel electrophoresis was used to monitor HEWL-gDNA binding. 20 µM HEWL was mixed with 10—200 µM gDNA. The increase in MRE, decrease in absorbance, and decrease in Trp fluorescence of the supernatants was monitored upon HEWL-gDNA mixing.

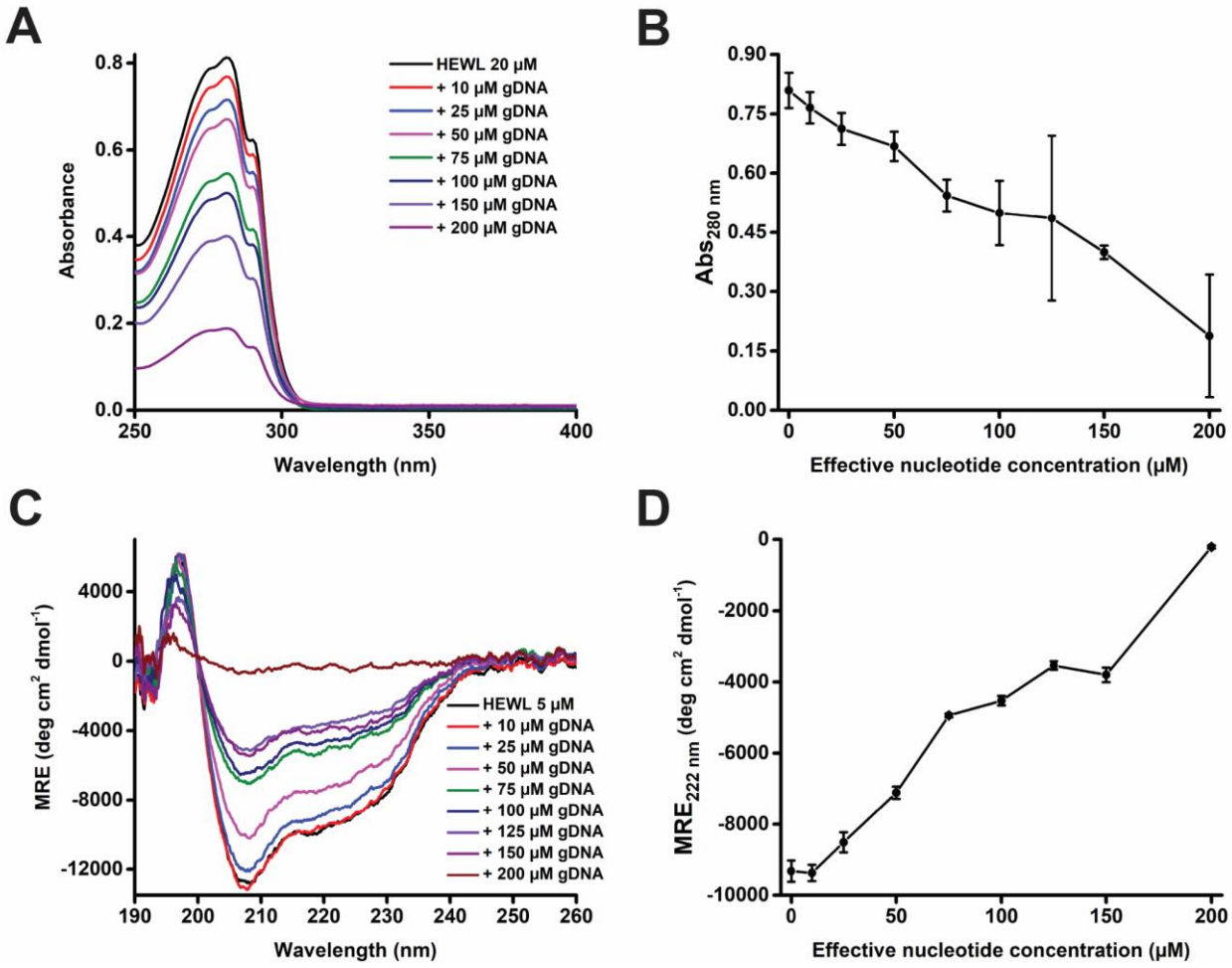
Substantial absorbance in the near UV region by aromatic amino acids was observed for 20 µM HEWL. Even the presence of low gDNA concentration (10 µM) caused a decrease in HEWL absorbance (Figure 6.2A). The decrease in HEWL absorbance was attributed to the precipitation of nucleoprotein complex upon HEWL-gDNA binding. The precipitation of nucleoprotein complex led to the decrease in supernatant HEWL concentration. The $Abs_{280\ nm}$ of HEWL-gDNA mixture supernatants decreased further in the presence of higher gDNA concentration (Figure 6.2B). HEWL absorbance maximum was at 281 nm and did not shift in the HEWL-gDNA mixture supernatants.

The observed HEWL circular dichroism was due to the presence of secondary structure elements in the protein. The far UV CD spectra of 5 µM HEWL was collected in the 190—260 nm region (Figure 6.2C). The ellipticity of HEWL-gDNA mixture supernatants increased substantially at higher DNA concentrations (25—200 µM). The increase in mean residual ellipticity was attributed

to reduced HEWL concentration in supernatants due to nucleoprotein complex precipitation with increasing gDNA concentration (Figure 6.2D). Due to almost complete precipitation of HEWL when mixed with 200 µM gDNA, there was absence of circular dichroism in the HEWL + 200 µM gDNA mixture supernatant.

Decrease in 10 µM HEWL Trp fluorescence was used to monitor nucleoprotein complex precipitation with increasing gDNA concentration. HEWL Trp were selectively excited at 295 nm and the fluorescence spectra of HEWL-gDNA mixture supernatants recorded in the presence of increasing gDNA concentration (Figure 6.2E). The decrease in fluorescence was observed even with low amount of gDNA and further decrease was observed at higher DNA concentrations. Negligible fluorescence was obtained from the HEWL + 200 µM gDNA sample. Hence the decrease in HEWL Trp fluorescence area under the curve upon mixing with 10—200 µM gDNA was used to monitor gDNA-HEWL binding indirectly (Figure 6.2F). The fluorescence emission maximum was at ~331 nm and shift was not observed upon HEWL-gDNA binding.

Precipitation of gDNA in HEWL-gDNA mixture supernatants was confirmed after agarose gel electrophoresis (Figure 6.2G). The electrophoresis of 25 µM only gDNA control samples and DNA-protein mixture supernatants expected to contain equivalent gDNA amount was performed. Absence of DNA in HEWL-gDNA mixtures confirmed the precipitation of gDNA by HEWL.



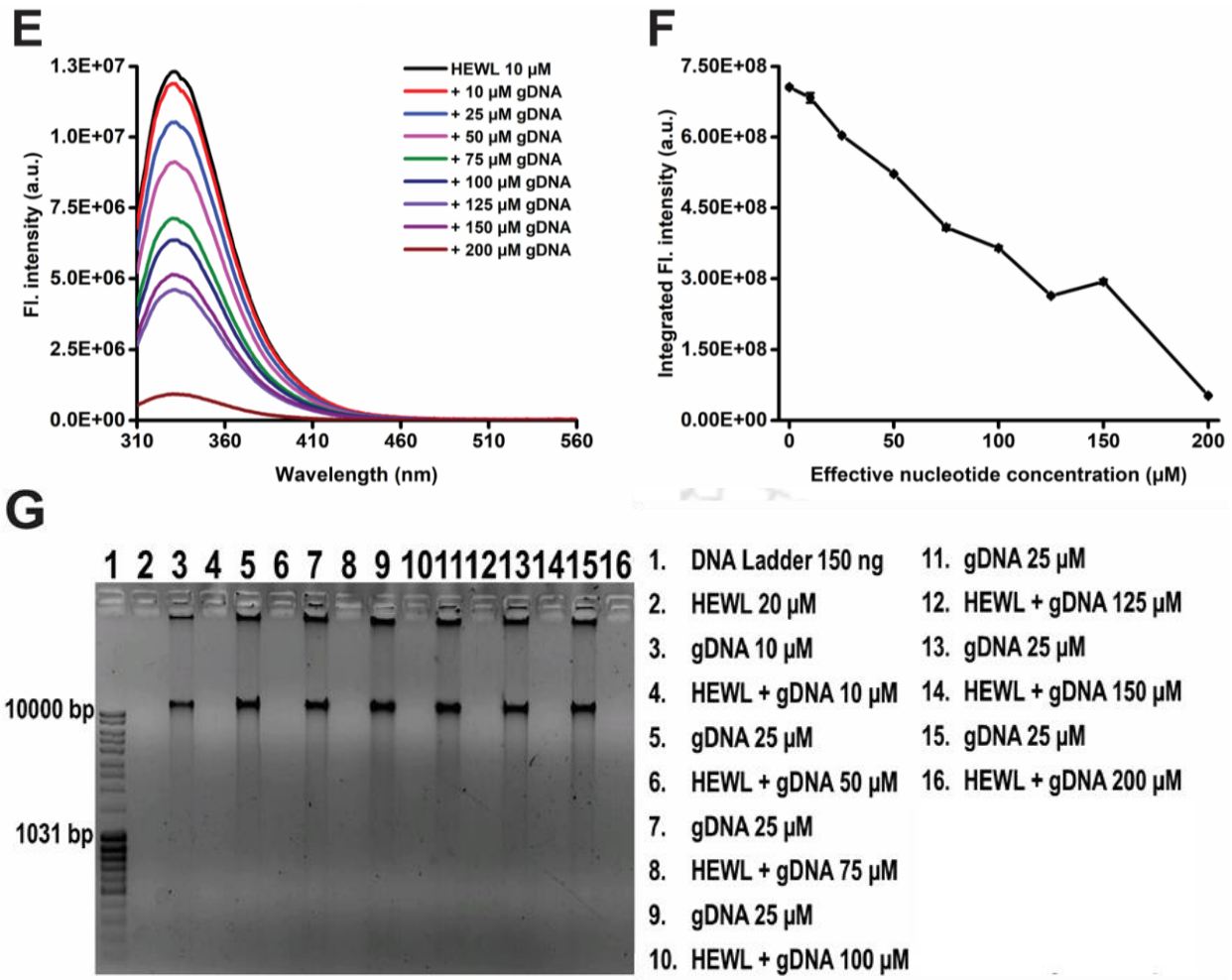


Figure 6.2: Monitoring the binding of HEWL with gDNA using HEWL absorbance and fluorescence. The binding of gDNA and HEWL was monitored using HEWL absorbance, HEWL Trp fluorescence, and mean residual ellipticity with increasing gDNA concentration and the data plotted. To 20 µM HEWL, increasing concentrations of gDNA (10—200 µM) were added. Sample preparation and assay were performed in 50 mM NaCl. DNA and HEWL mixing was done at 4 °C for 15 minutes at 120 rpm. After gDNA-HEWL mixing, centrifugation at 12000 rpm was done for 15 minutes and supernatant collected. HEWL absorbance was measured (Figure 6.2A) and decrease in absorbance at 280 nm (Figure 6.2B) was plotted. MRE was measured between 190—260 nm (Figure 6.2C) and increase in $MRE_{222\ nm}$ (Figure 6.2D) was used to characterize DNA-HEWL binding. Additionally, Trp fluorescence was monitored (Figure 6.2E) and decrease in integrated fluorescence intensity (Figure 6.2F) was plotted. 1 % agarose gel electrophoresis of the HEWL-gDNA mixture supernatants (Figure 6.2G) was used to visualize the absence of DNA.

The decrease in HEWL absorbance, decrease in HEWL Trp integrated fluorescence intensity, and increase in MRE_{222} were used to monitor HEWL-gDNA binding (Table 6.2). HEWL absorbance at 280 nm decreased from 0.809 to 0.188 in the supernatant of HEWL-gDNA mixture containing 200 µM gDNA. HEWL $MRE_{222\ nm}$ values increased substantially upon mixing with higher gDNA concentrations. Trp integrated fluorescence intensity decreased by 92.6 % in the presence of 200 µM gDNA.

gDNA concentration (μM)	$Abs_{280\ nm}$	STDEV	$MRE_{222\ nm}$	STDEV	$Area_{295\ nm}$	STDEV
0	0.809	0.0447	-9322	295	7.06E+08	2.31E+06
10	0.765	0.0394	-9374	228	6.84E+08	1.03E+07
25	0.712	0.0403	-8512	284	6.03E+08	8.96E+05
50	0.668	0.0372	-7111	172	5.22E+08	1.66E+06
75	0.544	0.0405	-4939	52	4.08E+08	5.29E+06
100	0.499	0.0819	-4522	134	3.65E+08	4.51E+06
125	0.486	0.2087	-3536	117	2.64E+08	1.54E+06
150	0.399	0.0171	-3801	204	2.94E+08	4.81E+06
200	0.188	0.1550	-210	41	5.25E+07	7.20E+05

Table 6.2: Monitoring the binding of HEWL with gDNA using HEWL absorbance and fluorescence. The binding of gDNA by HEWL was monitored using HEWL absorbance, Trp fluorescence, and MRE at 222 nm (deg cm^2 $dmol^{-1}$) with increasing gDNA concentration. The data was tabulated along with the standard deviation.

6.1.3 Monitoring the binding of PRM with gDNA using DNA absorbance

The absorbance of calf thymus gDNA (120 μM effective nucleotide concentration) was monitored upon mixing with PRM (Figure 6.3A). Efficient condensation and subsequent precipitation of gDNA by PRM was evident from the decrease in DNA absorbance at 260 nm (Figure 6.3B) and increased scattering at higher PRM concentrations (Figure 6.3C). The absorbance of PRM-gDNA mixtures was recorded after centrifugation and the scattering of PRM-gDNA mixtures was measured before centrifugation. The absorbance of PRM-gDNA mixture supernatant containing 0.6 μM PRM did not decrease hinting towards the solubility of nucleoprotamine complex at low PRM concentration. The absorbance at 260 nm decreased and the scattering at 350 nm increased substantially in the PRM-gDNA mixture containing 7.2 μM PRM (Figure 6.3D). The decrease in observed absorbance was attributed to a reduction in the amount of DNA in supernatant because of nucleoprotamine complex precipitation, and was also evident in the CD spectra.

In the absence of PRM, gDNA CD spectra had the features of a B-form DNA with a positive band at ~275 nm and a negative band at ~245 nm. The shape of the CD spectra remained the same in PRM-DNA mixture supernatants, however the ellipticity reduced with increasing PRM concentration (Figure 6.3E—F). The similarity in shape suggested that the possible adoption of alternate DNA conformation upon gDNA-PRM binding was not detected by CD spectroscopy. The reduced ellipticity was attributed to decrease in total nucleotide concentration in supernatant because of nucleoprotamine complex precipitation. Additionally, changes in PRM secondary structure could not be resolved from the far UV CD spectra due to precipitation of PRM in the working concentration range.

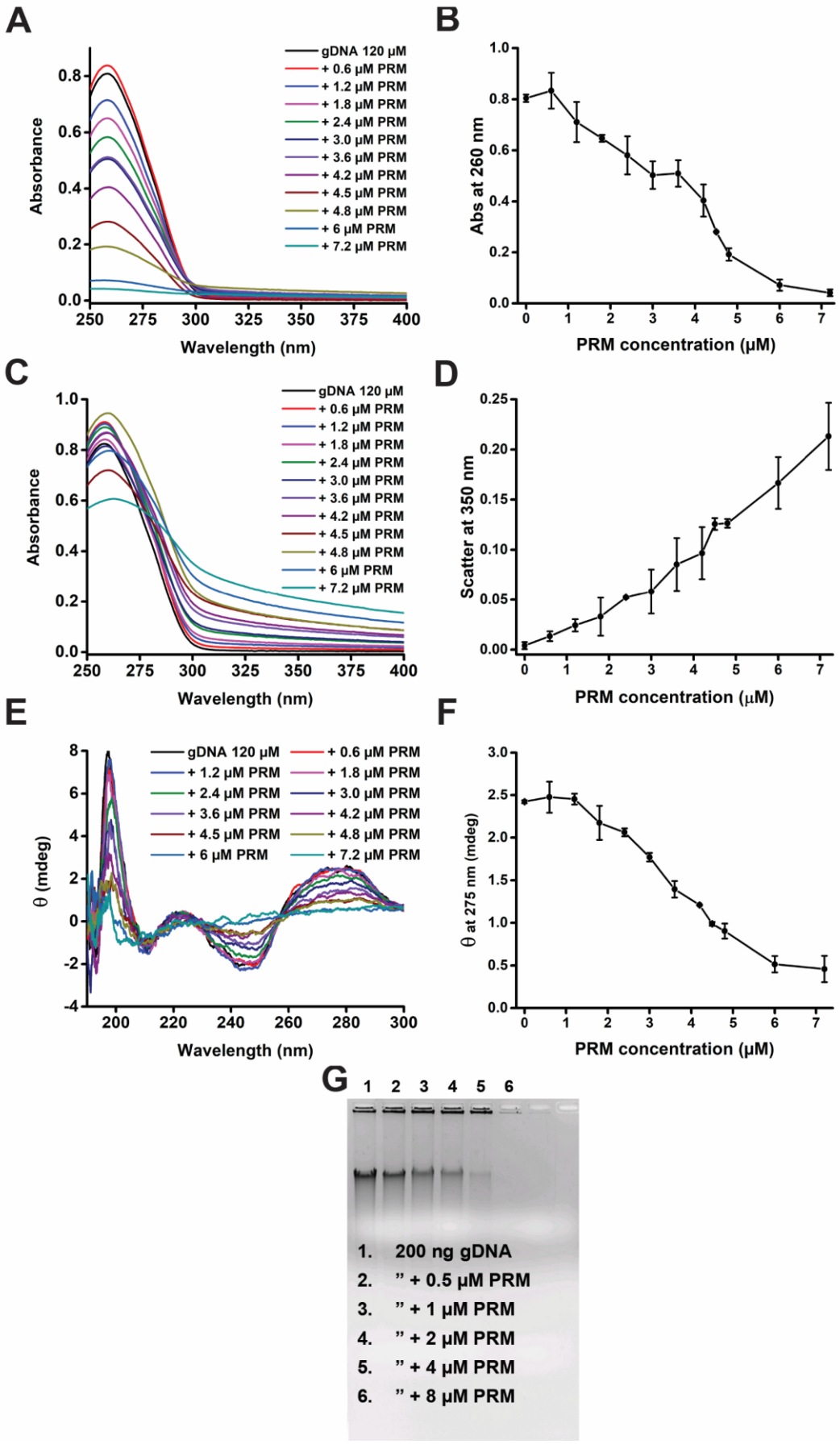


Figure 6.3: Monitoring the binding of PRM with gDNA using DNA absorbance. The binding of PRM with 120 µM gDNA was monitored using DNA absorbance, scattering of

nucleoprotamine complex, and CD spectroscopy. To a fixed concentration of gDNA (equivalent to 120 µM effective nucleotide concentration), increasing concentrations of PRM (0.6—7.2 µM) were added. Sample preparation and assay were performed in 50 mM NaCl. PRM and DNA mixing was done at 4 °C for 15 minutes at 120 rpm. After PRM-gDNA mixing, centrifugation at 12000 rpm was done for 15 minutes. Absorbance was monitored between 250—400 nm (Figure 6.3A) and absorbance at 260 nm (Figure 6.3B) was plotted. The scattering of nucleoprotein complex was collected for PRM-gDNA mixtures before centrifugation (Figure 6.3C) and plotted at 350 nm (Figure 6.3D). The scattering of PRM-gDNA mixtures at 350 nm was monitored before centrifugation. CD spectra was measured between 190—300 nm (Figure 6.1E) and ellipticity at 275 nm (Figure 6.1F) of the supernatant was used to characterize DNA-PRM binding. Electrophoretic mobility shift assay (Figure 6.3G) of the supernatants were used to characterize DNA-PRM binding.

PRM binding to gDNA caused a shift in the electrophoretic mobility of sheared DNA at low protein concentrations (Figure 6.3G). Lane 1 contained 200 ng sheared gDNA (control without PRM) with two prominent bands corresponding to higher mobility sheared DNA and low mobility unsheared gDNA. The amount of the sheared DNA decreased with increasing PRM concentration. This decrease coincided with an increase in low electrophoretic mobility band intensity for the PRM-DNA mixture containing 0.5 and 1 µM PRM. This increase was attributed to the formation of nucleoprotamine complex. At moderate PRM concentrations (2 and 4 µM), the sheared DNA further decreased in intensity. At higher PRM concentrations, a complete precipitation of gDNA happened as evidenced by the absence of DNA in the DNA + 8 µM PRM sample.

The decrease in DNA absorbance at 260 nm, decrease in ellipticity at 275 nm, and increased scattering by nucleoprotein complex at 350 nm was used to monitor gDNA-PRM binding (Table 6.3). An increasing amount of PRM was mixed with a fixed amount of gDNA. In the presence of 7.2 µM PRM, $Abs_{260\ nm}$ of the gDNA-PRM mixture supernatant decreased from 0.804 to 0.041. Since B-form DNA is low in chirality, the ellipticity values were low and in the presence of 7.2 µM PRM, $\theta_{275\ nm}$ decreased marginally from 2.422 mdeg to 0.459 mdeg. In the absence of PRM, almost no scattering was observed from only gDNA sample. Nucleoprotamine complex formation upon gDNA-PRM mixing caused an intense increase in scattering from 0.004 to 0.213 in the presence of 7.2 µM PRM. The scattering at 350 nm increased with each higher PRM concentration used.

PRM concentration (μM)	$Abs_{260\ nm}$	STDEV	$Scatter_{350\ nm}$	STDEV	$\theta_{275\ nm}$	STDEV
0	0.804	0.0146	0.004	0.0035	2.422	0.0170
0.6	0.833	0.0699	0.013	0.0049	2.477	0.1847
1.2	0.711	0.0788	0.024	0.0062	2.453	0.0641
1.8	0.647	0.0137	0.033	0.0190	2.174	0.2001
2.4	0.581	0.0743	0.052	0.0008	2.062	0.0470
3	0.503	0.0537	0.058	0.0220	1.771	0.0499
3.6	0.510	0.0516	0.085	0.0264	1.395	0.0966
4.2	0.404	0.0632	0.097	0.0261	1.209	0.0065
4.5	0.281	0.0005	0.126	0.0058	0.989	0.0223
4.8	0.191	0.0243	0.126	0.0044	0.904	0.0879
6	0.072	0.0221	0.167	0.0258	0.514	0.0969
7.2	0.041	0.0130	0.213	0.0334	0.459	0.1549

Table 6.3: Monitoring the binding of PRM with gDNA using DNA absorbance. The binding of PRM with 120 μM gDNA was monitored using DNA absorbance, ellipticity at 275 nm (reported in mdeg), and scattering at 350 nm upon mixing with increasing PRM concentrations.

6.1.4 Monitoring the binding of PRM with gDNA using PRM ProCharTS

PRM ProCharTS absorbance was monitored at multiple wavelengths after PRM binding with calf thymus gDNA. A significant decrease in 90 μM PRM ProCharTS was observed (Figure 6.4A) even at low concentration of DNA (effective nucleotide to amino acid ratio = 0.069). PRM ProCharTS decreased with increasing gDNA concentration and PRM ProCharTS at 250 nm reduced by 63.2 % in the presence of 700 μM gDNA (Figure 6.4B). Upon gDNA-PRM binding, a decrease in PRM ProCharTS was expected due to formation and subsequent precipitation of nucleoprotamine complex reducing the PRM concentration in PRM-gDNA mixture supernatants. Shape of the absorbance spectra suggested the absence of DNA in PRM-gDNA mixture supernatants. This implied that in the concentration range used, PRM caused complete condensation and precipitation of the DNA. Precipitation of the nucleoprotein complex was also observed when the binding of HEWL with gDNA was monitored using HEWL absorbance and fluorescence.

In the absence of DNA, PRM CD spectra had the characteristic features of a disordered protein with a minimum at ~200 nm (Figure 6.4C). Moderate amounts of DNA addition did not change the PRM CD spectra. However, the CD spectra of PRM samples mixed with higher gDNA concentrations were similar in shape but significantly reduced in intensity. The similarity of shape suggested that a possible change in PRM secondary structure upon DNA binding was not detected by the PRM far UV CD spectra. The circular dichroism decreased due to removal of PRM in the supernatants of PRM-gDNA mixtures because of nucleoprotamine complex precipitation. The characteristic near UV CD peaks expected for B-form gDNA were absent from PRM-DNA mixture supernatants. Absence of gDNA in PRM-DNA mixture supernatants was also visualized after agarose gel electrophoresis of the supernatants (Figure 6.4E).

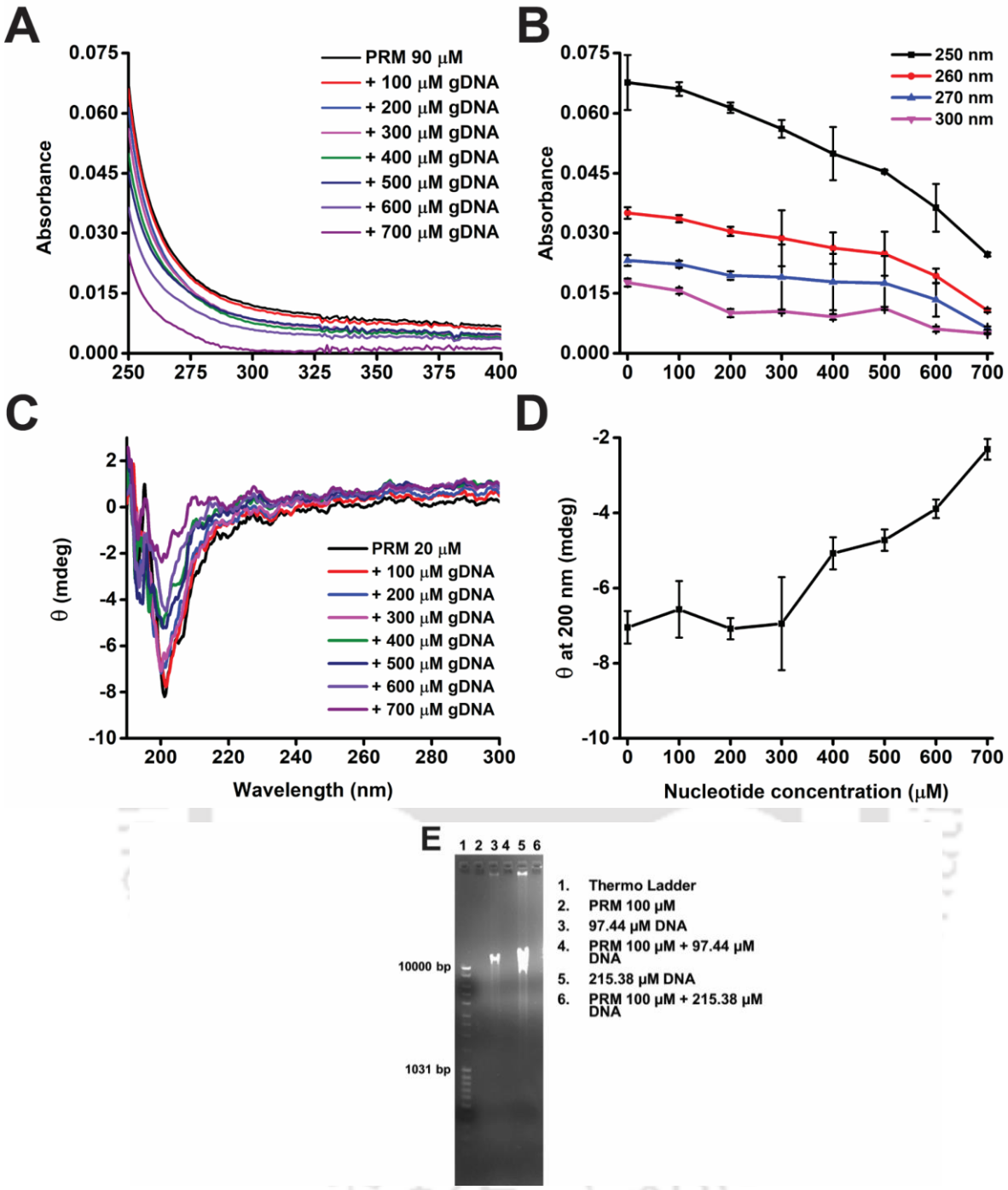


Figure 6.4: Monitoring the binding of PRM with gDNA using PRM ProCharTS. To a fixed concentration of PRM, an increasing amount of gDNA was added. Sample preparation and assay were performed in 50 mM NaCl. PRM and DNA mixing was done at 4 °C for 15 minutes at 120 rpm. After PRM-gDNA mixing, the samples were centrifuged at 12000 rpm for 15 minutes and the supernatant used for absorbance (Figure 6.4 A—B) and CD spectroscopy (Figure 6.4C—D) measurements. Agarose gel electrophoresis (Figure 6.4E) was used to ascertain the absence of DNA in PRM-gDNA mixture supernatants.

The decrease in PRM ProCharTS absorbance and increase in PRM ellipticity at 200 nm was used to monitor PRM-gDNA binding (Table 6.4). The absorbance of the PRM-gDNA mixture supernatants at 260 nm decreased from 0.035 to 0.011 in the presence of 700 µM gDNA. PRM

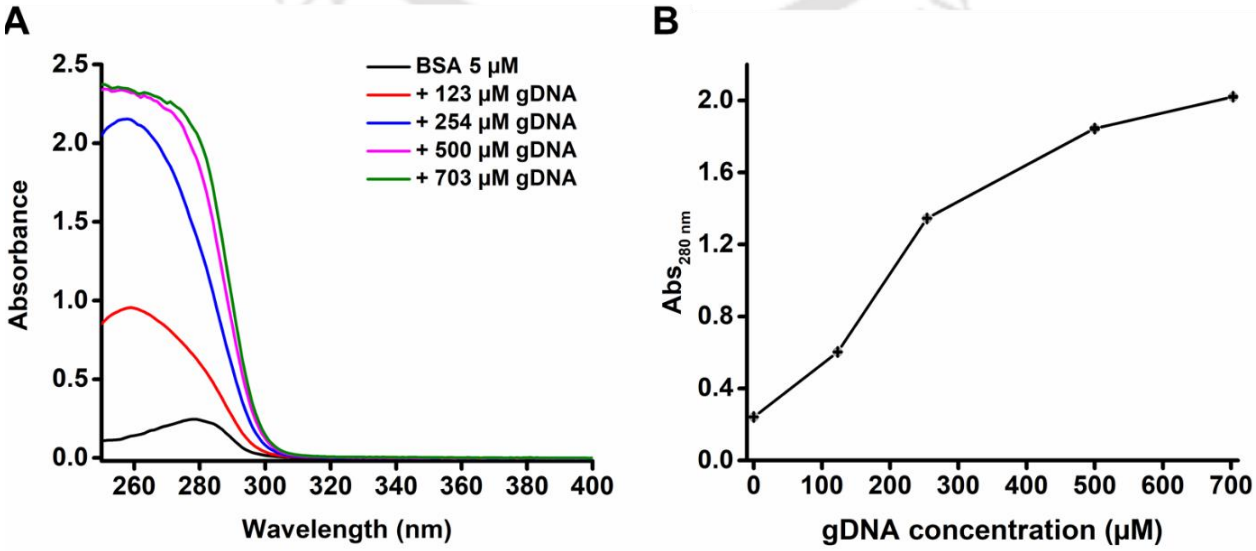
ellipticity increased from -7.05 to -2.31 mdeg. PRM ProCharTS at 250, 260, 270, 300 nm and ellipticity at 200 nm in the presence of increasing gDNA concentrations were tabulated and used to monitor the binding of PRM with gDNA.

gDNA concentration (μM)	$Abs_{250\ nm}$	STDEV	$Abs_{260\ nm}$	STDEV	$Abs_{270\ nm}$	STDEV	$Abs_{300\ nm}$	STDEV	$\theta_{200\ nm}$	STDEV
0	0.068	0.007	0.035	0.001	0.023	0.001	0.018	0.001	-7.05	0.432
100	0.066	0.002	0.034	0.001	0.022	0.001	0.016	0.001	-6.57	0.755
200	0.061	0.001	0.030	0.001	0.019	0.001	0.010	0.001	-7.08	0.284
300	0.056	0.002	0.029	0.007	0.019	0.008	0.011	0.000	-6.95	1.237
400	0.050	0.007	0.026	0.004	0.018	0.007	0.009	0.001	-5.08	0.427
500	0.045	0.000	0.025	0.005	0.018	0.007	0.011	0.000	-4.72	0.284
600	0.036	0.006	0.019	0.002	0.013	0.004	0.006	0.001	-3.89	0.245
700	0.025	0.000	0.011	0.000	0.006	0.000	0.005	0.000	-2.31	0.275

Table 6.4: Monitoring the binding of PRM with gDNA using ProCharTS absorbance. The binding of gDNA and PRM was monitored using PRM ProCharTS absorbance at 3 wavelengths and ellipticity at 200 nm (reported in mdeg) with increasing gDNA concentration.

6.1.5 Elucidating the absence of DNA condensation by BSA using condensation assay

DNA condensation assay was performed and BSA absorbance monitored to study the possible condensation of gDNA by BSA. The absorbance maximum of BSA was at 278 nm (Figure 6.5A) due to intense absorbance by the two tryptophans and other aromatic amino acids in BSA. Upon mixing of 5 μM BSA with gDNA, the absence of nucleoprotein complex precipitation was seen. The absorbance maximum of BSA + 123 μM gDNA was at 259 nm showing the presence of DNA in the BSA-DNA mixture supernatant (Figure 6.5A). The absorbance of BSA + gDNA mixture supernatants increased at higher gDNA concentration (Figure 6.5B) which was in contrast to the observations made earlier with PRM and HEWL. The protein PRM is a known DNA condensing agent and HEWL-DNA binding has been reported earlier. Earlier reports of BSA not being a DNA condensing agent was corroborated by the presence of gDNA in BSA-gDNA mixture supernatants observed after agarose gel electrophoresis as well (Figure 6.5C—D).



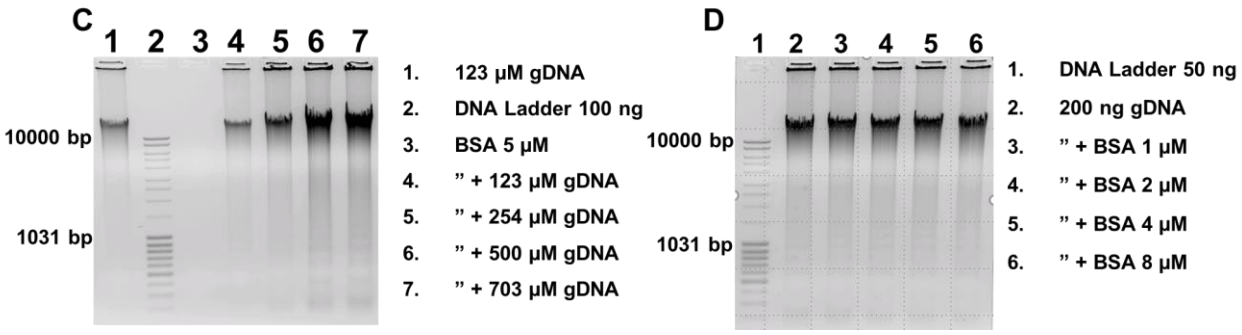


Figure 6.5: The absence of gDNA condensation by BSA monitored using BSA absorbance and agarose gel electrophoresis. DNA condensation assay was performed with 5 µM BSA and increasing gDNA concentration. Sample preparation and assay were performed in 50 mM NaCl. The absorbance spectra of BSA-gDNA mixture supernatants was measured in the 250—400 nm region (Figure 6.5A). The absorbance at 280 nm with increasing gDNA concentration was reported as a scatter plot (Figure 6.5B). 1 % agarose gel electrophoresis was used to monitor the presence of gDNA in BSA-gDNA mixture supernatants. To a fixed concentration of BSA, increasing amounts of DNA were mixed, and the presence of DNA in supernatant was monitored after gel electrophoresis (Figure 6.5C). The absence of DNA condensation was additionally confirmed by mixing fixed concentration of gDNA with increasing amounts of BSA (Figure 6.5D).

6.2 Conclusions

1. The binding of HEWL and PRM with calf thymus genomic DNA was monitored spectrophotometrically using decrease in gDNA absorbance, HEWL absorbance, HEWL Trp fluorescence, or PRM ProCharTS absorbance.
2. The condensation and subsequent precipitation of gDNA upon PRM binding was additionally confirmed using agarose gel electrophoresis.
3. HEWL and PRM contain an abundance of arginines and have a net positive charge at neutral pH, which allows the proteins to non-specifically bind with DNA.
4. The inability of BSA to cause DNA condensation was illustrated upon showing the presence of DNA in gDNA-BSA mixture supernatants.
5. The condensation of gDNA by PRM could not be rigorously distinguished from subsequent precipitation of the nucleoprotamine complex.
6. DNA condensation assays were used to monitor the decrease in DNA absorbance or protein absorbance/fluorescence of gDNA-protein mixture supernatants.
7. The decrease in PRM ProCharTS absorbance was used to illustrate the application of novel protein charge transfer spectra to monitor DNA-protein binding.

CHAPTER 7

DISCUSSION AND FUTURE DIRECTIONS

The phenomena of protein charge transfer absorbance and luminescence were observed in α_3C, α_3W, and PRM. Based on existing computational studies, the presence of SS-CT and PBS-CT transitions were predicted in α_3C and α_3W, whereas PBS-CT was predicted in PRM. Linearity of the absorbance and luminescence increase with concentration suggested that monomeric α_3C, α_3W, and PRM was the source of the observed ProCharTS.

The effect of solvent parameters on α_3C, α_3W, and PRM ProCharTS absorbance, luminescence, and secondary structure was monitored. The structural perturbations induced by extreme changes in solvent pH correlated well with the decrease in α_3C and α_3W ProCharTS at pH 13. α_3W ProCharTS absorbance decrease was more substantial at pH 13 compared to α_3C absorbance. Although α_3W is a point mutant of α_3C, the residual helical content at pH 13 was also greater in α_3W compared to α_3C. Thus single Trp substitution caused a significant difference in α_3W ProCharTS and secondary structure at alkaline pH. In the presence of salt (250 mM NaCl), α_3C, α_3W, and PRM amino acid sidechain charges would be significantly screened and a decrease in ProCharTS was expected and observed. Understanding the effect of solvent pH and the presence of salt on ProCharTS luminescence was less straightforward.

The sensitivity of ProCharTS to structural transitions was employed by Ansari *et al* to monitor structure gain in human c-Myc PEST fragment.[35] The potential application of ProCharTS to monitor structural transitions becomes more crucial for the class of intrinsically disordered proteins as they are rich in charged amino acid residues and contain little/no aromatic amino acid residues. Absence of the aromatic amino acid intrinsic chromophore limits the UV-Visible spectroscopic experiments that can be performed on IDPs.

The deviation of NATA fluorescence intensity decay from single to a sum of two exponentials in the presence of α_3C CT states was observed. The simultaneous contribution of Trp fluorescence and ProCharTS luminescence was resolved in α_3W intensity decay upon excitation at 295 nm by maximum entropy method analysis. ProCharTS luminescence component was significantly smaller in amplitude compared to Trp fluorescence. The significant difference in amplitude highlighted that due to low intensity, only a few reports of ProCharTS luminescence have been made.

Chemical denaturant-induced HSA, α_3W, and α_3C unfolding was characterized using ProCharTS and the conventional methods of CD spectroscopy, α_3W and HSA Trp fluorescence, α_3C-Dansyl and HSA-Dansyl fluorescence. Decrease in HSA, α_3W, and α_3C ProCharTS absorbance was observed upon protein unfolding. The absorbance decrease was attributed to the reduction in charged amino acid sidechain contacts upon unfolding. Decrease in charged amino acid sidechain contacts would significantly reduce the extent of PBS-CT and SS-CT in unfolded protein. Upon HSA, α_3W, and α_3C unfolding, ProCharTS absorbance at 325 nm decreased by 64, 76, and 72 % in the 6 M GdnHCl sample. HSA ProCharTS luminescence intensity decreased upon protein unfolding with the novel intrinsic luminescence being used for the first time to monitor protein unfolding.

HSA, α_3W, and α_3C are structured proteins containing 3 alpha-helices. The ellipticity recorded at multiple points in the protein unfolding pathways were used to perform secondary structure analysis. Upon protein unfolding, predominantly α-helix containing proteins HSA, α_3W, and α_3C became random coils. The alpha-helical content of folded HSA was in agreement with previous reports of HSA secondary structure obtained by CD and FTIR analysis.[96,97]

Nonlinear regression analysis was used to characterize HSA, α_3W, and α_3C unfolding pathway quantitatively. The denaturation midpoints estimated using ProCharTS were lower compared to the conventional techniques employed. The midpoints obtained suggested that disruption of charged amino acid contacts preceded the other markers of protein unfolding.

Additionally, ProCharTS absorbance was employed to detect a previously reported molten globule-like intermediate in GdnHCl induced protein unfolding pathways. Ahmed *et al* reported this intermediate in the HSA unfolding pathway in the presence of ~1.8 M GdnHCl.[6] The presence of the intermediate was confirmed by increased native Trp fluorescence, increased ANS fluorescence, and similarity of secondary structure in sub-denaturing GdnHCl concentrations.[6] Protein stiffening observed in multiple proteins at sub-denaturing denaturant concentrations has been confirmed by the analysis of X-ray crystal structures.[99,100] Protein stiffening is defined as the reduced dynamics of protein atoms around their average position.[57] Acharya *et al* and Dunbar *et al* stated that protein stiffening was a result of H-bonding between the polar atoms in urea/GdnHCl and protein main chain/sidechain atoms.[99,100] Thus the possibility of increased charged amino acid sidechain proximity in molten globule-like states explains the higher HSA ProCharTS absorbance observed. Folding intermediates and molten globule-like states can potentially be discovered in the unfolding pathways of charge-rich proteins at subdenaturing GdnHCl using the demonstrated increase in ProCharTS.

The application of ProCharTS to quantitatively monitor protein unfolding was illustrated for the first time in this study. Denaturation midpoint analysis allowed a comparison of the various methods used to monitor HSA, α_3W, and α_3C unfolding, thus providing a framework on which further studies to monitor the unfolding of proteins rich in charged amino acid residues using ProCharTS could be performed. It is clear from the studies done that ProCharTS absorbance is sensitive to the disruption of charged amino acid sidechain contacts for multiple charge-rich proteins. However, computational studies including MD simulation and TDDFT calculations are required to understand the exact molecular contacts getting disrupted. The approach of combining spectroscopic and computational measurements led to the discovery and characterization of the chromophore involved in the observed novel phenomenon by multiple researchers.[9,10] Computational studies will provide glimpses into the disruption of spatial proximities among multiple charged residues during protein unfolding. This will allow uncovering the fractional contribution of the type of charged amino acid sidechain contact (sequence adjacent or distally separated) disruption to the decrease in ProCharTS absorbance.

Previous reports have illustrated how DNA condensation by PRM cannot be distinguished from the subsequent precipitation of the nucleoprotein complex.[70] To a fixed concentration of gDNA, increasing amounts of HEWL/PRM were added and the decrease in DNA absorbance was used to monitor nucleoprotein complex formation and precipitation. In the presence of 0.5 µM HEWL and 0.6 µM PRM, the absorbance of the gDNA-protein mixture supernatant remained the same as only DNA sample. In the presence of moderate and high protein concentrations (1—4 µM HEWL and 1.2—7.2 µM PRM) the supernatant absorbance decreased substantially with increasing protein concentration. An increase in gDNA-protein mixture scattering at 350 nm was observed in the presence of 0.5—3.5 µM HEWL and 0.6—7.2 µM PRM. Monitoring the decrease in DNA absorbance was used previously by Gupta *et al* to observe DNA condensation by PRM at low protein concentration, and subsequent precipitation at higher PRM concentrations.[19]

The decrease in gDNA absorbance, PRM ProCharTS absorbance, HEWL absorbance, and HEWL Trp fluorescence was attributed to the decrease in gDNA/protein concentration in DNA-protein mixture supernatants due to the precipitation of nucleoprotein complex. Agarose gel electrophoresis was used to confirm the absence of DNA in gDNA-HEWL/PRM mixture supernatants after mixing fixed concentration of protein with increasing amounts of gDNA.

The previously reported inability of BSA to cause DNA condensation was shown when DNA absorbance was detected in gDNA-BSA mixture supernatants. DNA was also visualized in gDNA-BSA mixture supernatants after gel electrophoresis of the samples.

The DNA condensing behaviour of many biomolecules including protamines, spermine^{4+},[101] spermidine^{3+} (two physiologically occurring polycations),[102] polylysines,[103] histone protein H1,[104] and H5[105] have been previously reported. Apart from understanding the physiologically important process of DNA condensation and nucleoprotamine complex formation, other applications of DNA condensing agents include their possible use as transfection agents. It is clear that DNA condensation is a fascinating process which also has potential therapeutic applications. The use of ProCharTS thus provides a previously unknown spectroscopic window to monitor DNA condensation and precipitation *in-vitro* in further studies. PRM ProCharTS was used as a LABEL-FREE probe to monitor DNA condensation and precipitation in the present study. This is thus the first report of the application of novel Protein Charge Transfer absorbance spectra to detect DNA-protein binding. The principle used to detect PRM-gDNA binding can be applied to monitor the condensation of DNA by known and novel DNA condensing proteins in further studies.

The list of proteins which shows ProCharTS absorbance and luminescence is expanding rapidly with potential applications of the technique being discovered by multiple research groups.[45,82] It is expected that in further studies the process of protein unfolding and DNA binding of charge-rich proteins will be extensively characterized using this novel technique. This previously unknown spectroscopic window will provide additional opportunities of studying these important biomolecular processes *in-vitro.* In the coming decade, significant insights into the processes of protein unfolding and DNA binding is expected to be obtained from the technique of ProCharTS.

APPENDIX

1) Polyacrylamide gel electrophoresis:

Composition of resolving and stacking gels:

Components	Resolving gel (15 %)	Stacking gel (5 %)	Stock solutions
Acrylamide solution	5 mL	0.83 mL	29 % acrylamide, 1 % bisacrylamide
Resolving gel buffer	2.5 mL	-	1.5 M Tris.HCl (pH 8.8)
Stacking gel buffer	-	0.63 mL	0.5 M Tris.HCl (pH 6.8)
SDS solution	100 µL	50 µL	10 % (w/v) SDS
TEMED	5 µL	5 µL	undiluted
APS solution#	100 µL	50 µL	10 % (w/v) APS
H_20	2.3 mL	3.4 mL	
Total volume	10 mL	5 mL	

#: Freshly prepared

Coomassie staining solution:

1) 0.25 % (w/v) Brilliant Blue R

2) 40 % (v/v) methanol

3) 7 % (v/v) acetic acid

Destaining solution:

1) 30 % v/v Methanol

2) 10 % v/v glacial acetic acid

5 X gel loading buffer:

1) 0.05 % bromophenol blue

2) 40 % sucrose

3) 0.1 M EDTA (pH 8.0)

4) 0.5 % SDS.

5) 4.2 % β-mercaptoethanol

10 X Tris-glycine cathode buffer:

1) 30 g Tris base

2) 144 g Glycine

3) 10 g SDS

The three components were mixed in 1 L of H_2O to get 10 X running buffer of pH ~8.3 with pH adjustment not being required.

1 X Tris-Tricine cathode buffer

1) 0.1 M Tris

2) 0.1 M Tricine

3) 0.1 % SDS

5 X anode buffer

0.2 M Tris.Cl (pH adjusted to 8.9)

2) Amino acid sequences of commercially procured proteins used in the study

>P02768 Human Serum Albumin (HSA)

DAHKSEVAHR FKDLGEENFK ALVLIAFAQY LQQCPFEDHV
KLVNEVTEFA KTCVADESAE NCDKSLHTLF GDKLCTVATL
RETYGEMADC CAKQEPERNE CFLQHKDDNP NLPRLVRPEV
DVMCTAFHDN EETFLKKYLY EIARRHPYFY APELLFFAKR
YKAAFTECCQ AADKAACLLP KLDELRDEGK ASSAKQRLKC
ASLQKFGERA FKAWAVARLS QRFPKAEFAE VSKLVTDLTK
VHTECCHGDL LECADDRADL AKYICENQDS ISSKLKECCE
KPLLEKSHCI AEVENDEMPA DLPSLAADFV ESKDVCKNYA
EAKDVFLGMF LYEYARRHPD YSVVLLLRLA KTYETTLEKC
CAAADPHECY AKVFDEFKPL VEEPQNLIKQ NCELFEQLGE
YKFQNALLVR YTKKVPQVST PTLVEVSRNL GKVGSKCCKH
PEAKRMPCAE DYLSVVLNQL CVLHEKTPVS DRVTKCCTES
LVNRRPCFSA LEVDETYVPK EFNAETFTFH ADICTLSEKE
RQIKKQTALV ELVKHKPKAT KEQLKAVMDD FAAFVEKCCK
ADDKETCFAE EGKKLVAASQ AALGL

>P00698 Hen Egg White Lysozyme (HEWL)

KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT
QATNRNTDGS TDYGILQINS RWWCNDGRTP GSRNLCNIPC
SALLSSDITA SVNCAKKIVS DGNGMNAWVA WRNRCKGTDV
QAWIRGCRL

```
>P02769 Bovine Serum Albumin (BSA)
DTHKSEIAHR      FKDLGEEHFK      GLVLIAFSQY      LQQCPFDEHV
KLVNELTEFA      KTCVADESHA      GCEKSLHTLF      GDELCKVASL
RETYGDMADC      CEKQEPERNE      CFLSHKDDSP      DLPKLKPDPN
TLCDEFKADE      KKFWGKYLYE      IARRHPYFYA      PELLYYANKY
NGVFQECCQA      EDKGACLLPK      IETMREKVLA      SSARQRLRCA
SIQKFGERAL      KAWSVARLSQ      KFPKAEFVEV      TKLVTDLTKV
HKECCHGDLL      ECADDRADLA      KYICDNQDTI      SSKLKECCDK
PLLEKSHCIA      EVEKDAIPEN      LPPLTADFAE      DKDVCKNYQE
AKDAFLGSFL      YEYSRRHPEY      AVSVLLRLAK      EYEATLEECC
AKDDPHACYS      TVFDKLKHLV      DEPQNLIKQN      CDQFEKLGEY
GFQNALIVRY      TRKVPQVSTP      TLVEVSRSLG      KVGTRCCTKP
ESERMPCTED      YLSLILNRLC      VLHEKTPVSE      KVTKCCTESL
VNRRPCFSAL      TPDETYVPKA      FDEKLFTFHA      DICTLPDTEK
QIKKQTALVE      LLKHKPKATE      EQLKTVMENF      VAFVDKCCAA
DDKEACFAVE      GPKLVVSTQT      ALA
```

Figure A1: Amino acid sequence of proteins used in the study. Charged amino acids are highlighted in red and aromatic amino acids are shown in light blue.

Protein	No of Positively charged amino acids (% +ve charge)	No of Negatively charged amino acids (% -ve charge)	Total charge content (%)
HEWL	18 (13.9)	9 (7.0)	20.9
HSA	99 (16.9)	98 (16.8)	33.7
BSA	99 (16.9)	99 (16.9)	33.8

Figure A2: Charged amino acid content of HEWL, HSA, and BSA

3) Equations used in nonlinear regression analysis:

The following inbuilt functions in Origin 2016 software were used to perform the nonlinear regression to obtain protein unfolding denaturation midpoints:

1) Boltzmann function: $y = \frac{A_1 - A_2}{1 + e^{(x - x_0)/dx}} + A_2$ A1.1

2) Logistic function: $y = \frac{A_1 - A_2}{1 + (x/x_0)^p} + A_2$ A1.2

3) Exponential decay function: $y = y_0 + A_1 e^{\frac{-x}{t_1}}$ A1.3

4) Dose-response function: $y = A1 + \frac{A2 - A1}{1 + 10^{(LOGX_0 - X)p}}$ A1.4

4) Monitoring protein purity by Polyacrylamide Gel Electrophoresis:

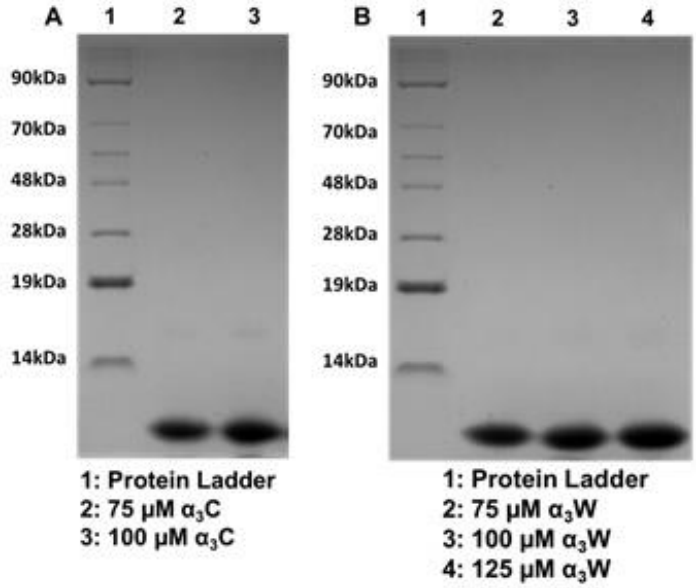


Figure A3: Assessment of protein purity by gel electrophoresis. (A) 15 % SDS-PAGE of α_3C. (B) 15 % SDS-PAGE of α_3W.

4) Monitoring protein purity by MALDI-TOF spectroscopy:

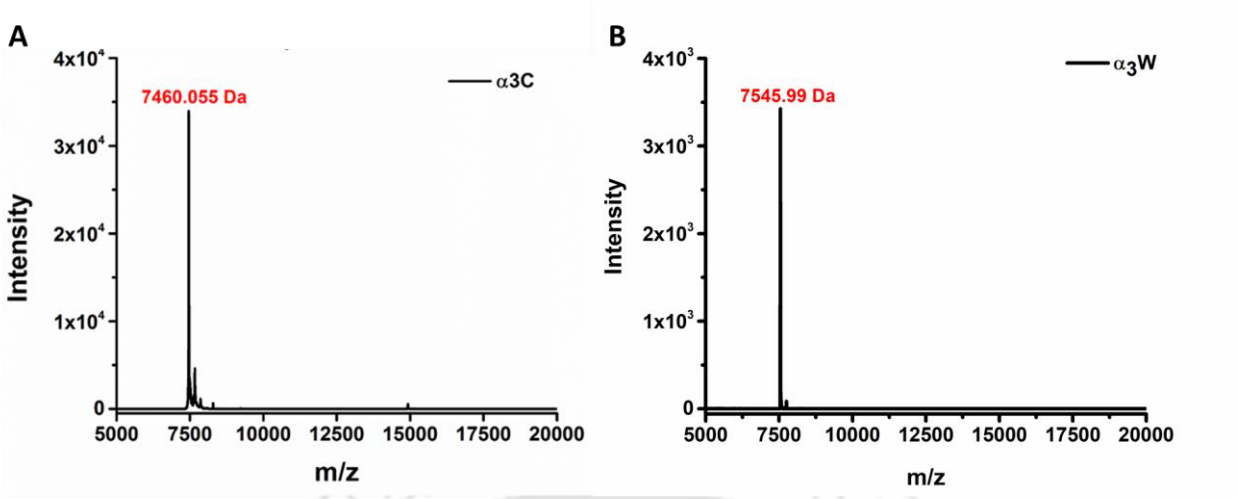


Figure A4: Assessment of purified protein purity by Mass spectroscopy. (A) A sharp peak at 7460.055 Da was observed. Calculated Mass of α_3C = 7460 Da. (B) The mass obtained was 7545.99 Da which is very close to the mass calculated from the sequence of α_3W = 7543 Da.

LIST OF PUBLICATIONS AND CONFERENCES

PUBLICATIONS:

1. Kumar, A.; Alom, S.E.; Ahari, D.; Priyadarshi, A.; Ansari, M.Z.; Swaminathan, R. Role of Charged Amino Acids in Sullying the Fluorescence of Tryptophan or Conjugated Dansyl Probe in Monomeric Proteins. *Biochemistry* **61**, 339–353 (2022)
2. Kumar, A.; Ahari, D.; Priyadarshi, A.; Ansari, M.Z.; Swaminathan, R. Weak Intrinsic Luminescence in Monomeric Proteins Arising from Charge Recombination. *J. Phys. Chem. B.* **124**, 2731–2746 (2020)
3. Ansari, M.Z.; Kumar, A.; Ahari, D.; Priyadarshi, A.; Padmavathi, L.; Bhandari, R.; Swaminathan, R. Protein Charge Transfer Absorption Spectra: An Intrinsic Probe to Monitor Structural and Oligomeric Transitions in Proteins. *Faraday Discuss.* **207**, 91-113 (2018)
4. Priyadarshi, A. and Swaminathan, R. The application of Protein Charge Transfer Spectra to monitor Protein unfolding. (*Under peer review*)
5. Priyadarshi, A. and Swaminathan, R. Monitoring DNA-protein binding using Protein Charge Transfer Spectra. (*Manuscript under preparation*)

LIST OF CONFERENCE PAPERS:

1. Biophysical Society Annual Meeting 2022.
   Title: Monitoring Protein Unfolding and DNA Binding using Protein Charge Transfer Spectra
   Priyadarshi, A. and Swaminathan, R.
   Biophysical Journal, 121.3 (2022): 185a.

2. Biophysical Society Annual Meeting 2020.
   Title: Unravelling the Origin of Multi-Exponential Fluorescence Intensity Decay of Tryptophan in Protein.
   Kumar, A.; Alom, S.E.; Ahari, D.; Priyadarshi, A.; Ansari, M.Z.; Swaminathan, R.
   Biophysical Journal, 118(3) (2020): 469a.

3. Biophysical Society Annual Meeting 2020.
   Title: Weak Intrinsic Luminescence in Monomeric Proteins Arising from Charge Recombination.
   Kumar, A.; Ahari, D.; Priyadarshi, A.; Ansari, M.Z.; Swaminathan, R.
   Biophysical Journal, 118(3) (2020): 468a.

4. Biophysical Society Annual Meeting 2018.
   Title: Protein Charge Transfer Absorption Spectra: An Intrinsic Probe to Monitor Structural and Oligomeric Transitions in Proteins.
   Ansari, M.Z.; Kumar, A.; Ahari, D.; Priyadarshi, A.; Padmavathi, L.; Bhandari, R.; Swaminathan, R.
   Biophysical Journal, 114(3) (2018): 586a

CONFERENCES ATTENDED:

1. Light-matter Interactions from scratch: Theory and Experiments at the Border with Biology, 2021, ICTP, Italy
2. Research Conclave 2019, IIT Guwahati
3. National Workshop on Fluorescence 2018, JNU
4. FCS Workshop 2017, IIT Guwahati
5. INPEC 2017, Bose Institute, Kolkata
6. Optics Within Life Sciences (OWLS) 2016, Tata Institute of Fundamental Research

REFERENCES

1. Wüthrich, K. Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem.* **265**, 22059–22062 (1990).
2. Haris, P. I. & Chapman, D. Does Fourier-transform infrared spectroscopy provide useful information on protein structures? *Trends Biochem. Sci.* **17**, 328–333 (1992).
3. Rygula, A. *et al.* Raman spectroscopy of proteins: a review. *J. Raman Spectrosc.* **44**, 1061–1076 (2013).
4. Thomas, P. J., Qu, B.H. & Pedersen, P. L. Defective protein folding as a basis of human disease. *Trends Biochem. Sci.* **20**, 456–459 (1995).
5. von Hippel, P. H. From 'simple' DNA-protein interactions to the macromolecular machines of gene expression. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 79–105 (2007).
6. Ahmad, B., Ahmed, M. Z., Haq, S. K. & Khan, R. H. Guanidine hydrochloride denaturation of human serum albumin originates by local unfolding of some stable loops in domain III. *Biochim. Biophys. Acta - Proteins Proteomics* **1750**, 93–102 (2005).
7. Yu, S. S. & Li, H. J. Helix–coil transition and conformational studies of protamine–DNA complexes. *Biopolymers* **12**, 2777–2788 (1973).
8. Cantor, C. & Schimmel, P. Techniques for the study of biological structure and function. (*W.H. Freeman*, 1980).
9. Prasad, S. *et al.* Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. *Chem. Sci.* **8**, 5416–5433 (2017).
10. Pinotsi, D. *et al.* Proton Transfer and Structure-Specific Fluorescence in Hydrogen Bond-Rich Protein Structures. *J. Am. Chem. Soc.* **138**, 3046–3057 (2016).
11. Niyangoda, C., Miti, T., Breydo, L., Uversky, V. & Muschol, M. Carbonyl-based blue autofluorescence of proteins and amino acids. *PLoS One* **12**, 1–15 (2017).
12. Homchaudhuri, L. & Swaminathan, R. Novel Absorption and Fluorescence Characteristics of L-Lysine. *Chem. Lett.* **30**, 844–845 (2001).
13. 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).
14. 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).
15. 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).
16. 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).
17. Balhorn, R. The protamine family of sperm nuclear proteins. *Genome Biol.* **8**, 227 (2007).
18. Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat. Protoc.* **2**, 1849–1861 (2007).
19. Gupta, S., Tiwari, N. & Munde, M. A Comprehensive Biophysical Analysis of the Effect of DNA Binding Drugs on Protamine-induced DNA Condensation. *Sci. Rep.* **9**, 1–11 (2019).
20. Steinrauf, L. K., Shiuan, D., Yang, W. & Chiang, M. Y. Lysozyme Association with Nucleic

Acids. *Biochem. Biophys. Res. Commun.* **266**, 366–370 (1999).

21. Plaxco, K. W. & Dobson, C. M. Time-resolved biophysical methods in the study of protein folding. *Curr. Opin. Struct. Biol.* **6**, 630–636 (1996).
22. Borgia, A., Williams, P. M. & Clarke, J. Single-molecule studies of protein folding. *Annu. Rev. Biochem.* **77**, 101–125 (2008).
23. Joseph R. Lakowicz. Principles of Fluorescence Spectroscopy. (*Springer US*, 2006).
24. Chang, R. Physical Chemistry for the Biosciences. (*University Science Books*, 2005).
25. Wetlaufer, D. B. Ultraviolet spectra Of Proteins and Amino Acids. (*Advances in Protein Chemistry*, 1963).
26. Demchenko, A. P. Ultraviolet Spectroscopy of Proteins. (*Springer Berlin*, 1987).
27. Du, R., Yang, D., Jiang, G., Song, Y. & Yin, X. An Approach for In Situ Rapid Detection of Deep-Sea Aromatic Amino Acids Using Laser-Induced Fluorescence. *Sensors* **20**, (2020).
28. Bernofsky, C. & Wanda, S. Y. Formation of reduced nicotinamide adenine dinucleotide peroxide. *J. Biol. Chem.* **257**, 6809–6817 (1982).
29. Falk, J. E. Porphyrins and metalloporphyrins their general, physical and coordination chemistry, and laboratory methods. (*Elsevier Publishing Company*, 1964).
30. de la Lande, A., Gillet, N., Chen, S. & Salahub, D. R. Progress and challenges in simulating and understanding electron transfer in proteins. *Arch. Biochem. Biophys.* **582**, 28–41 (2015).
31. Gray, H. B. & Winkler, J. R. Electron transfer in proteins. *Annu. Rev. Biochem.* **65**, 537–561 (1996).
32. Tennent, D. L. & McMillin, D. R. A detailed analysis of the charge-transfer bands of a blue copper protein. Studies of the nickel(II), manganese(II), and cobalt(II) derivatives of azurin. *J. Am. Chem. Soc.* **101**, 2307–2311 (1979).
33. Giese, B., Graber, M. & Cordes, M. Electron transfer in peptides and proteins. *Curr. Opin. Chem. Biol.* **12**, 755–759 (2008).
34. Stoner-Ma, D. *et al.* Observation of Excited-State Proton Transfer in Green Fluorescent Protein using Ultrafast Vibrational Spectroscopy. *J. Am. Chem. Soc.* **127**, 2864–2865 (2005).
35. 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).
36. 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).
37. Shukla, A. *et al.* A novel UV laser-induced visible blue radiation from protein crystals and aggregates: Scattering artifacts or fluorescence transitions of peptide electrons delocalized through hydrogen bonding? *Arch. Biochem. Biophys.* **428**, 144–153 (2004).
38. Chen, X. *et al.* Prevalent intrinsic emission from nonaromatic amino acids and poly(amino acids). *Sci. China Chem.* **61**, 351–359 (2018).
39. Chan, F. T. S. *et al.* Protein amyloids develop an intrinsic fluorescence signature during aggregation. *Analyst* **138**, 2156–2162 (2013).
40. 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).

41. Pansieri, J. *et al.* Ultraviolet–visible–near-infrared optical properties of amyloid fibrils shed light on amyloidogenesis. *Nat. Photonics* **13**, 473–479 (2019).
42. Guptasarma, P. Solution-state characteristics of the ultraviolet A-induced visible fluorescence from proteins. *Arch. Biochem. Biophys.* **478**, 127–129 (2008).
43. Jong, K. H. *et al.* Low energy optical excitations as an indicator of structural changes initiated at the termini of amyloid proteins. *Phys. Chem. Chem. Phys.* **21**, 23931–23942 (2019).
44. Tikhonova, T. N. *et al.* Dissection of the deep-blue autofluorescence changes accompanying amyloid fibrillation. *Arch. Biochem. Biophys.* **651**, 13–20 (2018).
45. Ziaunys, M., Sneideris, T. & Smirnovas, V. Exploring the potential of deep-blue autofluorescence for monitoring amyloid fibril formation and dissociation. *PeerJ* **7**:e7554 (2019).
46. Nelson, D. L. Principles of biochemistry. (*W.H. Freeman*, 2005).
47. Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).
48. Dill, K. A. Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155 (1990).
49. Schiffer, C. A. & Dötsch, V. The role of protein-solvent interactions in protein unfolding. *Curr. Opin. Biotechnol.* **7**, 428–432 (1996).
50. Santra, M. K., Banerjee, A., Rahaman, O. & Panda, D. Unfolding pathways of human serum albumin: Evidence for sequential unfolding and folding of its three domains. *Int. J. Biol. Macromol.* **37**, 200–204 (2005).
51. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
52. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* **37**, 215–246 (2008).
53. Prakash, S. & Matouschek, A. Protein unfolding in the cell. *Trends Biochem. Sci.* **29**, 593–600 (2004).
54. Jones, C. M. *et al.* Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl. Acad. Sci.* **90**, 11860–11864 (1993).
55. Sabelko, J., Ervin, J. & Gruebele, M. Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci.* **96**, 6031–6036 (1999).
56. Roder, H. & Shastry, M. C. R. Methods for exploring early events in protein folding. *Curr. Opin. Struct. Biol.* **9**, 620–626 (1999).
57. Kumar, R., Prabhu, N. P., Yadaiah, M. & Bhuyan, A. K. Protein stiffening and entropic stabilization in the subdenaturing limit of guanidine hydrochloride. *Biophys. J.* **87**, 2656–2662 (2004).
58. Serrano, A. L., Waegele, M. M. & Gai, F. Spectroscopic studies of protein folding: linear and nonlinear methods. *Protein Sci.* **21**, 157–170 (2012).
59. Rischel, C. & Poulsen, F. M. Modification of a specific tyrosine enables tracing of the end-to-end distance during apomyoglobin folding. *FEBS Lett.* **374**, 105–109 (1995).
60. Oliva, R. Protamines and male infertility. *Hum. Reprod. Update* **12**, 417–435 (2006).
61. Lewis, J. D., Song, Y., De Jong, M. E., Bagha, S. M. & Ausió, J. A walk though vertebrate

and invertebrate protamines. *Chromosoma* **111**, 473–482 (2003).

62. Hud, N. V, Milanovich, F. P. & Balhorn, R. Evidence of Novel Secondary Structure in DNA-Bound Protamine Is Revealed by Raman Spectroscopy. *Biochemistry* **33**, 7528–7535 (1994).

63. Roque, A., Ponte, I. & Suau, P. Secondary structure of protamine in sperm nuclei: an infrared spectroscopy study. *BMC Struct. Biol.* **11**, 14 (2011).

64. Cai, Y.H. & Huang, H. Advances in the study of protein-DNA interaction. *Amino Acids* **43**, 1141–1146 (2012).

65. Dey, B. *et al.* DNA-protein interactions: methods for detection and analysis. *Mol. Cell. Biochem.* **365**, 279–299 (2012).

66. Orlando, V. Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem. Sci.* **25**, 99–104 (2000).

67. Ladbury, J. E. & Chowdhry, B. Z. Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. *Chem. Biol.* **3**, 791–801 (1996).

68. Srivastava, V. K. & Yadav, R. Isothermal titration calorimetry. (*Academic Press*, 2019).

69. Kypr, J., Kejnovská, I., Bednářová, K. & Vorlíčková, M. Circular Dichroism Spectroscopy of Nucleic Acids. *Comprehensive Chiroptical Spectroscopy* **17**, 575–586 (2012).

70. Bloomfield, V. A. DNA condensation by multivalent cations. *Biopolymers* **44**, 269–282 (1997).

71. Schweizer, T., Kubach, H. & Koch, T. Investigations to characterize the interactions of light radiation, engine operating media and fluorescence tracers for the use of qualitative light-induced fluorescence in engine systems. *Automot. Engine Technol.* **6**, 275–287 (2021).

72. C., E. L. & Bevington, P. R. Data Reduction and Error Analysis for the Physical Sciences. *Journal of the American Statistical Association* **67** 249 (1972).

73. Livesey, A. K. & Brochon, J. C. Analyzing the Distribution of Decay Constants in Pulse-Fluorimetry Using the Maximum Entropy Method. *Biophys. J.* **52**, 693–706 (1987).

74. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2006).

75. Glover, S. D. *et al.* Photochemical Tyrosine Oxidation in the Structurally Well-Defined α_3Y Protein: Proton-Coupled Electron Transfer and a Long-Lived Tyrosine Radical. *J. Am. Chem. Soc.* **136**, 14039–14051 (2014).

76. Brewer, L. R., Corzett, M. & Balhorn, R. Protamine-induced condensation and decondensation of the same DNA molecule. *Science* **286**, 120–123 (1999).

77. Herskovits, T. T. & Brahms, J. Structural investigations on DNA·protamine complexes. *Biopolymers* **15**, 687–706 (1976).

78. Acharya, N., Mishra, P. & Jha, S. K. A dry molten globule-like intermediate during the base-induced unfolding of a multidomain protein. *Phys. Chem. Chem. Phys.* **19**, 30207–30216 (2017).

79. Pelosi, C. *et al.* Thermodynamic Evaluation of the Interactions between Anticancer Pt(II) Complexes and Model Proteins. *Molecules* **26**, 2376 (2021).

80. Uppuluri, K. B., Ayaz Ahmed, K. B., Jothi, A. & Veerappan, A. Spectrofluorimetric and molecular docking investigation on the interaction of 6-azauridine, a pyrimidine nucleoside

antimetabolite, with serum protein. *J. Mol. Liq.* **219**, 602–607 (2016).

81. Homchaudhuri, L., Kumar, S. & Swaminathan, R. Slow aggregation of lysozyme in alkaline pH monitored in real time employing the fluorescence anisotropy of covalently labelled dansyl probe. *FEBS Lett.* **580**, 2097–2101 (2006).

82. Ansari, M. Z. & Swaminathan, R. Structure and dynamics at N- and C-terminal regions of intrinsically disordered human c-Myc PEST reveal a pH-induced transition. *Proteins Struct. Funct. Bioinforma.* **88**, 889–909 (2020).

83. Louis-Jeune, C., Andrade-Navarro, M. A. & Perez-Iratxeta, C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins* **80**, 374–381 (2012).

84. Kawahara, K. & Tanford, C. Viscosity and Density of Aqueous Solutions of Urea and Guanidine Hydrochloride. *J. Biol. Chem.* **241**, 3228–3232 (1966).

85. Pletneva, E. V, Gray, H. B. & Winkler, J. R. Many Faces of the Unfolded State: Conformational Heterogeneity in Denatured Yeast Cytochrome c. *J. Mol. Biol.* **345**, 855–867 (2005).

86. Gupta, S., Tiwari, N. & Munde, M. A Comprehensive Biophysical Analysis of the Effect of DNA Binding Drugs on Protamine-induced DNA Condensation. *Sci. Rep.* **9**, 5891 (2019).

87. Tommos, C., Valentine, K. G., Martínez-Rivera, M. C., Liang, L. & Moorman, V. R. Reversible Phenol Oxidation and Reduction in the Structurally Well-Defined 2-Mercaptophenol-α_3C Protein. *Biochemistry* **52**, 1409–1418 (2013).

88. Dai, Q.H. *et al.* Structure of a de Novo Designed Protein Model of Radical Enzymes. *J. Am. Chem. Soc.* **124**, 10952–10953 (2002).

89. 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).

90. Toma, A. C., De Frutos, M., Livolant, F. & Raspaud, E. DNA condensed by protamine: A 'short' or 'long' polycation behavior. *Biomacromolecules* **10**, 2129–2134 (2009).

91. Brouwer, A. M. Standards for photoluminescence quantum yield measurements in solution (IUPAC technical report). *Pure Appl. Chem.* **83**, 2213–2228 (2011).

92. Kiefhaber, T., Schmid, F. X., Willaert, K., Engelborghs, Y. & Chaffotte, A. Structure of a rapidly formed intermediate in ribonuclease T1 folding. *Protein Sci.* **1**, 1162–1172 (1992).

93. Matsuo, K., Sakurada, Y., Yonehara, R., Kataoka, M. & Gekko, K. Secondary-structure analysis of denatured proteins by vacuum-ultraviolet circular dichroism spectroscopy. *Biophys. J.* **92**, 4088–4096 (2007).

94. Lundberg, D., Carnerup, A. M., Schillén, K., Miguel, M. da G. & Lindman, B. Phase Behavior and Coassembly of DNA and Lysozyme in Dilute Aqueous Mixtures: A Model Investigation of DNA−Protein Interactions. *Langmuir* **26**, 2986–2988 (2010).

95. Lundberg, D. *et al.* Size and morphology of assemblies formed by DNA and lysozyme in dilute aqueous mixtures. *Phys. Chem. Chem. Phys.* **13**, 3082–3091 (2011).

96. Cheema, M. A. *et al.* Human serum albumin unfolding pathway upon drug binding: A thermodynamic and spectroscopic description. *J. Chem. Thermodyn.* **41**, 439–447 (2009).

97. Usoltsev, D., Sitnikova, V., Kajava, A. & Uspenskaya, M. Systematic FTIR Spectroscopy Study of the Secondary Structure Changes in Human Serum Albumin under Various Denaturation Conditions. *Biomolecules* **9**, 359 (2019).

98. Jonas, H. *et al*. Conformational properties of cardiolipin-bound cytochrome c. *Proc. Natl. Acad. Sci.* **109**, 125–130 (2012).

99. Pike, A. C. W. & Acharya, K. R. A structural basis for the interaction of urea with lysozyme. *Protein Sci.* **3**, 706–710 (1994).

100. Dunbar, J., Yennawar, H. P., Banerjee, S., Luo, J. & Farber, G. K. The effect of denaturants on protein structure. *Protein Sci.* **6**, 1727–1733 (1997).

101. Chattoraj, D. K., Gosule, L. C. & Schellman, J. A. DNA condensation with polyamines: II. Electron microscopic studies. *J. Mol. Biol.* **121**, 327–337 (1978).

102. GOSULE, L. C. & SCHELLMAN, J. A. Compact form of DNA induced by spermidine. *Nature* **259**, 333–335 (1976).

103. Laemmli, U. K. Characterization of DNA condensates induced by poly(ethylene oxide) and polylysine. *Proc. Natl. Acad. Sci.* **72**, 4288–4292 (1975).

104. W., H. M. & David, C. R. Structure of histone H1-DNA complex: Effect of histone H1 on DNA condensation. *Proc. Natl. Acad. Sci.* **74**, 4852–4856 (1977).

105. Garcia-Ramírez, M. & Subirana, J. A. Condensation of DNA by basic protiens does not depend on protein composition. *Biopolymers* **34**, 285–292 (1994).