Analyzing Charge Transfer Spectra arising from non-aromatic amino acids in proteins, aggregating peptides and viral capsid assemblies

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

By

Shah Ekramul Alom

Roll No. 156106032

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati, Guwahati,
Assam- 781039, India

February 2023

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.

Shah Ekramul Alom

IIT Guwahati — Shah Ekramul Alom

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, entitled "***Analyzing Charge Transfer Spectra arising from non-aromatic amino acids in proteins, aggregating peptides and viral capsid assemblies",*** done by **Shah Ekramul Alom** 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 extremely delighted to be in this position in my life after all the hard work and sacrifices I have put through in the ten years of my B.Tech. and Ph.D. life at IIT Guwahati. I don't think I would have made it possible without the help of the people mentioned hereafter, and thus, I would like to express my sincere gratitude to all of them.

First and foremost, I want to express my deepest respect and sincere gratitude to my supervisor Prof. R. Swaminathan for giving me the opportunity to undertake Ph.D. in his research group and for having faith in me. His constant moral support, enthusiasm in research, and motivation during the tough times helped me immensely. I would like to thank him for his scientific insights into the domain of work related to this thesis. He has been a great mentor who encouraged me never to be afraid to try new research ideas, bring the best out of me, and make me a better researcher. I am indebted to him for guiding me in the right direction whenever I needed it the most.

I would also like to extend my sincere gratitude to the Doctoral Committee members, Prof. Biman Behari Mandal, Prof. Manish Kumar and Dr. Kanhaiya Pandey, for their encouragement, critical assessment of my Ph.D. work and valuable suggestions, which helped me a lot to troubleshoot the problems I had to face in my Ph.D.

Besides my Doctoral Committee, I would also like to extend my sincere thanks to Prof. Bhubaneswar Mandal for his encouragement and insightful suggestions, which helped me a lot in troubleshooting difficulties in my PhD. work. Also, I am extremely grateful to Prof. Bhubaneswar Mandal and his students, Altaf Hussain Kawa and Sourav Kalita, for synthesizing and providing me with the switch peptides used in my PhD. work.

I would like to thank Prof. Michael Blaber from Florida State University, USA, and Prof. Michael Nassal from the University of Freiburg, Germany, for providing the Symfoil 4P/PV2 and HBV Cp plasmid constructs, respectively.

My sincere thanks to the BSBE department for letting me use the DCIF facilities, and the staff members for being highly cooperative. Thanks to Central Instruments Facility for letting me use Mass spectrometry (MALDI), Transmission Electron Microscopy (TEM) and Atomic Force Microscopy (AFM) facilities. Thanks to IIT Guwahati for all the facilities and the Ministry of

Human Resource Development (MHRD), India, for the financial support. Thanks to Dihing Hostel administration and staff members, as the hostel life will hold a special place in my heart.

Next, I wish to express my sincere thanks to my lab seniors Tulsi Swain, Saumya Prasad and Shrutidhara Biswas, Mohd. Ziauddin Ansari, Amrendra Kumar, Dileep Ahari, Anurag Priyardarshi, and Sanjana for their enormous help and suggestion during my Ph.D. I would like to express my deep sense of gratitude to all my juniors Alka, Heeramoni, Nuzelu, Nayanjyoti, Garima, Subhajit, Abheek, Harikrishna, Simagka, Himanshi, Sahil, Kartik, Zunaid and Niteesh. I am especially thankful to Mohd. Ziauddin Ansari and Amrendra Kumar for helping me learn the techniques used in my Ph.D.

I would especially like to thank my friends Sattwik, Anuj, Jayant, Kunal, Mukesh, Jitender, Heeramoni, Alka, Abheek, Darshana, Nayanmoni, Bhoomika, Nimmi, Prateek for making my stay at IIT Guwahati a memorable one. Also, I would like to thank Simangka for being my lawn tennis partner and helping me to become a better player. Thanks to Kartik for helping me conduct some of the works from my Ph.D. thesis.

My Ph.D. endeavour could not have been completed without the never-ending love and endless support from my family. I would like to thank my parents, who have sacrificed their happiness for me and motivated me in every stage of my life. I am grateful to my late grandfather for his unending love and emotional support since my childhood. I would also like to thank my cousins and school friends for their visit to the institute campus and for never making me feel lonely during my stay at IIT Guwahati.

Although it is not possible to pen down the names of each individual people, I would like to thank each one of them for helping me and having trust in me.

Finally, my deepest thanks to my sister, Yasmin Nasima Parvin, who has been the most important person in my life in helping me get to this point. Her constant support, motivation and help during every difficult phase of my life is overwhelming and cannot be expressed in a few words. Her unconditional love has been a constant source of motivation throughout my entire life. She has truly been the best friend and the most valued person in my life.

Last but not least, I am thankful to Almighty God for his continuous blessing.

Shah Ekramul Alom *August 2022*

List of Abbreviations

ANS	8-Anilinonaphthalene-1-sulfonic acid
ThT	Thioflavin T
CD	Circular Dichroism
TCSPC	Time-Correlated Single Photon Counting
Dansyl	2-dimethyl amino naphthalene 6-sulfonyl
ProCharTS	Protein Charge Transfer Spectra
Aβ	Amyloid beta
SwP	Switch Peptide
DTT	1,4-Dithiothreitol
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TFA	Trifluoroacetic acid
UV-Vis	Ultraviolet-Visible
MALDI	Matrix-assisted laser desorption ionization
NATA	N-acetyl tryptophan amide
CT	Charge Transfer
ET	Electron Transfer
MEM	Maximum Entropy Method
NLLS	Non-Linear Least Square
E. Coli	Escherichia Coli
SEC	Size Exclusion Chromatography
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

DPA	9,10-Diphenylanthracene
IPTG	Isopropyl β-D-1-thiogalactopyranoside
HBV Cp	Hepatitis B Virus Core Protein
CpAMs	Core Protein Allosteric Modulators
DMF	Dimethyl Formamide
TDDFT	Time-Dependent Density Functional Theory
FRET	Förster Resonance Energy Transfer
Gdn.HCl	Guanidine Hydrochloride
1,5-IAEDANS	5-((((2-Iodoacetyl) amino) ethyl) amino) naphthalene-1-sulfonicacid
SEC	Size Exclusion Chromatography
DLS	Dynamic Light Scattering
TEM	Transmission Electron Microscopy

Thesis Abstract

Spectroscopic techniques, like electronic absorption or fluorescence, allow us to understand proteins and their structure-function relationship. Extrinsic probes (fluorescent dyes) and intrinsic chromophores (aromatic amino acids, prosthetic groups, peptide bonds) are routinely studied to decipher the knowledge about protein structure and dynamics. Till the 20th century, intrinsic UV-Visible absorption and fluorescence in apo-protein were thought to be arising from the conventional chromophores, i.e., aromatic amino acids like tryptophan (Trp), tyrosine (Tyr) and phenylalanine (Phe). Intrinsic chromophores like Trp yield information about their local fluctuations in the protein, surrounding the chromophore. However, the need to conjugate a protein with extrinsic probes to yield information about protein dynamics, can sometimes alter the protein structure. Thus, the importance for an intrinsic label-free alternatives showing similar UV-visible absorption and luminescence properties is a recent topic of interest. There have been reports of novel UV-Visible absorbance above 315 nm from various non-aromatic amino acids both in crystal and solution form. Similar to absorption, intrinsic deep-blue luminescence under UV light excitation, has been observed in charged monomeric proteins, crystals, aggregates and small molecules. However, the origin of such luminescence is still controversial. Electron Transfer among the charged moieties in these amino acids and proteins has been thought to be the most potential explanation behind these novel optical phenomena, however, it is still unclear. Although these novel phenomena have been often linked mainly to highly charged clusters of amino acids or oligomeric proteins and aggregated. Recently, our research group in collaboration with Venkatramani group, have shown the intrinsic UV-Visible absorption and fluorescence in a highly monomeric charged protein α_3C with no aromatic amino acids and proposed. Charge Transfer among various Lys-Glu, Lys-Lys and Glu-Glu contacts is expected to be the plausible reason behind the phenomenon. TDDFT calculations from various snapshots of MD simulations indicated the photoinduced electron transfer from electron donors (Glutamate anion or polypeptide backbone) to electron-deficient acceptors (Lysine cation or polypeptide backbone). Moreover, the subsequent charge recombination from the electron-deficient acceptors to the electron donors is proposed to be the reason behind the novel deep-blue luminescence. This phenomenon of photoinduced electron transfer and subsequent charge recombination has been termed **Protein Charge Transfer Spectra (ProCharTS)**. Most of the studies on ProCharTS absorbance and luminescence have been well characterized for proteins predominantly rich in Lysine and Glutamate amino acid residues.

However, no reports of *in vitro* studies of ProCharTS on proteins devoid of Lysine or Glutamate have been performed.

In Chapter 3, we report for the first time ProCharTS absorbance and luminescence in homo-polypeptides of Arginine and Aspartate and also in a highly charged protein Symfoil 4P/PV2 which has 126 amino acids with 30% charged content. Symfoil 4P/PV2 has an abundance of Arginine, Aspartate and Histidine residues, with no lysine or aromatic residues. The ProCharTS absorbance and luminescence arising from the Arginine, Aspartate and Histidine residues showed surprisingly similar features as that of Lysine and Glutamate and are highly conserved among all charged amino acids, peptides, and proteins. Absorbance in UV-Visible wavelength extending till 800 nm, deep-blue luminescence upon UV light excitation, accompanied by large Stokes shift, low quantum yield, excitation wavelength-dependent emission maxima, and fast luminescence lifetime decay are some of the novel characteristic features of ProCharTS. Moreover, the fast luminescence lifetime decay is characterized by the prominent contribution from the shorter luminescence lifetime (< 1ns), where the charge recombination phenomenon is quite fast. The lower quantum yield (Φ_{355} ~0.003-0.04) points to the fact that most of the photoexcitation (> 97%) relaxes via non-radiative transfer. The strong correlation between the slope of integrated luminescence intensity with the molar extinction coefficient and also with the product of extinction coefficient and quantum yield, signifies the major role of extinction coefficient in the ProCharTS luminescence. Thus, the origin of ProCharTS is evident as long as electron transfer takes place among the proximal charged species, irrespective of the kind of charged residues present in the system. However, the ProCharTS absorption and fluorescence output from Arginine, Aspartate and Histidine are relatively lower as compared to Lysine.

Chapter 4 brings forward the additional luminescence of charged amino acids, namely Arginine, Aspartate, and Histidine, on the indole fluorescence. The significant overlap of the absorbance and luminescence spectrum of these charge amino acids and the fluorescence from the indole group in NATA (a tryptophan analogue), sully the fluorescence of the indole group. Efforts have been made to address the question why NATA exhibits one exponential fluorescence decay, while Tryptophan shows two or more exponential fluorescence decay. The luminescence from these amino acids contaminating the fluorescence of the indole group in NATA is believed to contribute to the Trp-like fluorescence observed in proteins. Both the Non-linear least square (NLLS) and Maximum Entropy Method (MEM) analysis hint at the possible contribution of the emergence of the shorter fluorescence component from the indole

ring in Trp. Moreover, the charged residues were found to sully the indole fluorescence, irrespective of whether its fluorescence is getting quenched or not. Amino acids in clusters quench the NATA fluorescence more than that of proteins. In contrast, the close-spaced charged clusters in protein contribute more to the complexity of the fluorescent lifetime of the indole group.

In Chapters 5 and 6, the applications of ProCharTS on Aβ peptide aggregation and HBV Cp capsid assembly were performed. Aggregation of Aβ is basically an intermolecular event where the individual peptides stack together to form oligomers and subsequently form fibrils. Switch peptides (SwPs) of Aβ were synthesized here to systematically study the early events of Aβ aggregation. When pH changes from acidic (~pH 2) to neutral or slightly basic conditions (~pH 6.8-7.2), the O to N acyl migration in the molecular switch triggers the self-assembly. During $A\beta_{16\text{-}22}$ aggregation, the Phenylalanine residues in the hydrophobic core are stacked together. The initially random coiled peptides start to convert into β-sheet, and subsequently, the charged residues in the flanking ends come close, and ProCharTS is enhanced. A distance of 4.8 Å spaces the arrangement of each individual anti-parallel β-strand; thus, Lys/Lys or Lys/Glu interactions are expected during $A\beta_{16\text{-}22}$ aggregation. ProCharTS was able to detect the early events of SwP oligomerization and aggregation, which showed enhanced absorption and luminescence upon aggregate formation. Moreover, both 8-Anilinonaphthalene-1-sulfonic acid (ANS) and Thioflavin T (ThT) assays showed increased fluorescence upon aggregation and complimented the results obtained from both Circular Dichroism (CD) and ProCharTS. ProCharTS luminescence decay study revealed the characteristic 3 luminescence lifetime components with higher contribution (> 60%) from the shorter lifetime component (< 1ns). However, the mean luminescence lifetime decreased upon aggregation. The sensitivity of CD and ProCharTS absorbance was found to be higher than the rest of the techniques. Moreover, the kinetics of the early events of Aβ aggregation obtained from ProCharTS followed a sigmoidal profile and correlated well with CD data.

Similarly, Hepatitis B Virus core protein (HBV Cp) capsid assembly can be considered a systematically controlled aggregation where clusters of individual HBV core proteins arrange in an icosahedral symmetry enhancing molecular 3D contacts within charged species. HBV Cp dimer assembled under 50 mM HEPES plus 0.25-0.5 M NaCl (pH 7.5) and have been characterized by conventional methods: Circular Dichroism (CD), Dynamic Light Scattering (DLS), Atomic Force Microscopy (AFM), Transmission Electron Microscopy (TEM) and Size

Exclusion Chromatography (SEC). The kinetics of the HBV Cp 149 capsid assembly is studied by ProCharTS, Tryptophan fluorescence and 90° light scattering. The half-time of the capsid assembly ($t_{1/2}$) showed a faster rate kinetics (~ 2min) from 90° light scattering, whereas it was found to be marginally longer (3-5min) from ProCharTS and Tryptophan fluorescence study. The ProCharTS and intrinsic Tryptophan fluorescence, thus may be was able to pick up the relatively later stages of maturation and subsequent reorientation of the dimers in capsids. Next, the HBV Cp dimers were successfully labelled with both Dansyl Chloride and 1,5-IAEDANS. Compared to dimers, around 25% increase in fluorescence anisotropy for the extrinsic probes was observed upon capsid formation. However, the fluorescence anisotropy change for Tryptophan was significantly less upon capsid assembly, most probably because of the possibility of home FRET in the Trp-Trp pair among the nearby Tryptophan in dimer-dimer interface in the capsids.

Taken together, this study reveals ProCharTS phenomenon even in proteins devoid of Lysine residues and further enhances the applicability of ProCharTS in the field of protein aggregation and capsid assembly.

Table of Contents

Page No.

Chapter 2: Experimental Techniques used; Materials and Methods

Chapter 3: ProCharTS arising from Arginine, Aspartate amino acids observed in peptides & proteins which are devoid of Lysine residues

Chapter 4: Role of charged amino acids in sullying the fluorescence of Tryptophan in monomeric protein

Chapter 5: To monitor the early events of Amyloid beta (Aβ) Switch peptide aggregation in real-time using ProCharTS

Chapter 6: To monitor the Hepatitis B Virus Core protein (HBV Cp) capsid assembly in real-time using ProCharTS and extrinsic chromophores

Chapter 7: Thesis Summary and Future Directions

Chapter 1

Introduction and Review of Literature

Chapter 1

1.1. Introduction

Proteins are essential components in living systems, and understanding their role and structure-function relationship is a common topic in biophysics. Spectroscopic techniques like absorption and fluorescence allow us to study protein structural changes[1]. Optical features arising from extrinsic fluorescent dyes and intrinsic chromophores such as aromatic amino acids, prosthetic groups, and peptide bonds have been studied extensively in this regard[2]. Whether other intrinsic chromophores exist in proteins is a topic worth investigating. Spectroscopic features from such novel intrinsic probes can thus be used to detect biophysical processes in a non-invasive manner. They can further add new insights into the spectroscopic analysis in the UV-Visible range (200-800 nm). This problem statement broadly constitutes the topic of my project, and I am highlighting below the recent advancements in this field along with the objectives of my work.

1.2. Conventional chromophores in proteins

Molecules undergo electronic transitions such that their electrons are promoted from lower to higher molecular orbitals when excited by higher energy electromagnetic radiation. Most of these transitions lie in the UV (200-400 nm) and visible (400-800 nm) range of the electromagnetic spectrum. Although absorbance below 200 nm in the UV range from proteins is evident, electronic absorption from the water, which takes place below 170 nm, can contaminate them[3]. So, for studying biological macromolecules, we often restrict ourselves above 200 nm. The phenomena underlying the UV-Visible absorption by proteins and other macromolecules have been utilized as a conventional method in molecular biology since the early days[4]. Here, in this section, I tried to highlight briefly the most commonly involved chromophores employed in protein biology that are often used as a probe to detect structural and conformational changes in protein, along with the nature of the origin of such optical features.

1.2.1. Peptide bond

Valence electrons existing in any molecular orbital, whether being in single (σ-bonding orbitals), double/triple bonds (π bonding orbitals) or non-bonding orbitals (lone electron pair), can make electronic transitions from σ; π; n to σ*; π*[5] as shown in **Figure 1.1.** In the UV-

Visible spectrum, π-π* and n-π* transitions involving unsaturated centres are most prominent ones[6] whereas the σ-σ* and n-σ* transitions which are reflected in relatively lower wavelength due to larger energy gap, are generally less studied.

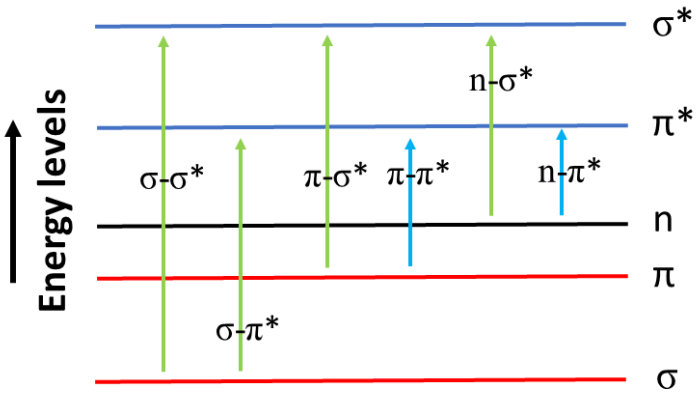


Figure 1.1: Molecular orbital of electrons with varying energy levels. The most prominent transitions (in blue) involve π-π* and n-π* transitions, and the less frequent transitions (in green) are also highlighted.

In proteins, the most common electronic transitions include the peptide bond and have characteristic absorbance in the far-UV region. Predominantly, the delocalized π valence electrons in the peptide bonds over the N, C, and O atoms and nonbonding n-orbital electrons near the O atom are involved. Two distinct peaks: one strong peak at 190 nm with extinction coefficient ε = 7000 $M^{-1}cm^{-1}$ (involving π- π* transitions) and another weaker intensity peak at about 210–220 nm with ε = 100 $M^{-1}cm^{-1}$ (involving n- π* transitions) characterizes the peptide bond[7,8]. In addition to both these peaks, a third peak at around 175 nm involving an n-σ* transition can also be observed.

1.2.2. Aromatic amino acids

Apart from the peptide backbone, which shows absorbance below the 220 nm region, the only known chromophores in an apoprotein with absorption features are the side chains of the aromatic amino acids Tryptophan (Trp), Tyrosine (Tyr) and Phenylalanine (Phe). Their optical properties display significant absorption at wavelengths longer than 230 nm (near-UV region) and are, in fact, the gold standard of protein quantification and estimation. Among the three aromatic amino acids, Trp shows the strongest absorption arising from the indole group present in the side chain. It is characteristic of two major peaks: at around 220 nm with ε = 36,000 $M^{-1}cm^{-1}$ and ε = 5,600 $M^{-1}cm^{-1}$ at 280 nm[9,10]. Tyr is one of the other aromatic residues having absorbance properties with ε = 1400 $M^{-1}cm^{-1}$ at 275 nm and with ε = 9000 $M^{-1}cm^{-1}$ at 222 nm.

The optical properties of the Tyr residues in proteins are highly pH dependent and the most among the aromatic amino acids[11]. Phe shows considerable absorbance in the UV region with two prominent absorption bands: one with ε = 200 $M^{-1}cm^{-1}$ at around 257 nm[12,13] and another band at around 205 nm with ε = 9600 $M^{-1}cm^{-1}$. In addition to the aromatic amino acids, the imidazole group in the side chain of Histidine, Cysteine and Methionine also display appreciable absorbance in the lower wavelength region between 185-220 nm and 230-240 nm range, respectively[13]. However, the disulphides (cystine) show absorption bands at longer wavelengths as compared to cysteine between 250-270 nm with ε of 300 $M^{-1}cm^{-1}$.

1.2.3. Charge Transfer complexes

Apart from the already mentioned chromophores as described above, there exists another category of chromophores which mainly rely upon Charge Transfer (CT) of electrons from donor to acceptor. Although such charge transfer complexes are relevant in various biomolecules like DNA, porphyrins, saccharides, lipids, vitamins and even in microorganisms, here only the prominent cases of Charge-Transfer observed in protein systems are highlighted. Various biological processes like photosynthesis in bacteria and plants[14] or that of electron transport chain in the respiratory pathway mainly rely on the formation of such Charge Transfer complexes. The commonly studied systems are two species of purple bacteria (Rhodobacter sphaeroides and Rhodopseudornonavsi ridis), where CT transitions occur between two bacterial chlorophylls. The most common and well-studied examples of biomolecules are metalloproteins and enzymes. Generally, metal-ligand complexes in metalloproteins show absorptions through charge transfer transitions[15]. It is generally categorized into two categories: ligand-to-metal charge-transfer (LMCT) bands (Eg. $KMnO_4$) or metal-to-ligand charge-transfer (MLCT) bands (For e.g. $[Ru(bpy)_3]Cl_2$)[16]. Such metal–protein complex (e.g., Azurin, stellacyanin, Xanthine oxidase, plastocyanin); heme-containing proteins (cytochrome c, b5, and myoglobin) along with many important coenzymes of proteins such as FAD, NADH and NAD^+ display strong absorption in UV-Visible region. In addition to metal-ligand complexes, transition metal series and inner transition metals belonging to the d and f electronic transitions alone can also display d-d transitions involve the electronic transitions in the filled and unfilled d-orbitals and show broad absorption bands in the visible region and most are brightly coloured[17]. The intensity of absorption is quite weak and is caused due to vibronic coupling, and are Laporte forbidden.

Charge Transfer (CT) mechanism has also been reported in peptides containing aromatic amino acids, Tyr and Trp[18], where the conversion of indolyl into phenoxy radical is mediated via electron transfer either through single step super-exchange[19] or by multistep hopping mechanism[20]. Moreover, the conformation of protein also plays an important role in CT. Electron transfer was found to be dependent upon the dipole moment, where faster rates are observed in the direction of the dipole moment of α-helix while slower rates in the opposite direction. Moreover, long-range electron transfer via the amide bonds through the α-helical peptides as intermediate relay junctions has also been reported[21,22]. Recently, the peptide bond and the side chain of charged amino acids have been reported to act like donor-acceptor pairs for the charge transfer phenomenon[23].

1.2.4. Non-conventional chromophores

Most of the chromophores mentioned above are either the chromophores present in apoproteins like peptides and aromatic amino acids or are metal complexes. Recently in the 21st century, there have been multiple reports of unusual UV-Visible absorbance spectra above 315 nm and sometimes extending up to 800 nm as observed in the aqueous solution of charged amino acid S/R-Lysine[24], L-Lysine-HCl[25] and poly-L-lysine-HCl[26]. Similar absorbance has been found in various amino acids[27,28] and other solutes[29], but the chromophore giving rise to such spectra was not known. In contrast to absorption, intrinsic luminescence from not only proteins but also from various biopolymers, nucleic acids as well as carbohydrates has been a recent topic of interest. There have been many reports of intrinsic fluorescence or deep-blue/blue luminescence arising from monomeric proteins[30], oligomeric proteins[31,32], crystals[33], and amyloid aggregates[34,35]. Similarly, dendrimers (PAMAM) and various small molecular structures[36], synthetic and natural polymers, which lacked any conventional fluorophores, also showed such intrinsic blue luminescence.

The chromophores responsible for these systems are mainly the C-O, N-H, C-N, C=O, and COOH groups and not from the indole or phenyl ring in aromatic amino acids, as shown in **Figure 1.2**. Hydrogen bonding and other weak interactions, Through-space conjugation (TSC) arising from n–σ*, n–π*, and π–π* orbital overlap have been proposed to be the origin of such luminescence. Proton transfer and electron delocalization are also thought to be the possible reason for such phenomenon. But whether this luminescence is due to the above-mentioned

hypothesis or any other phenomenon is still unresolved. Apart from fluorescence, phosphorescence and room temperature phosphorescence have also been reported[37,38].

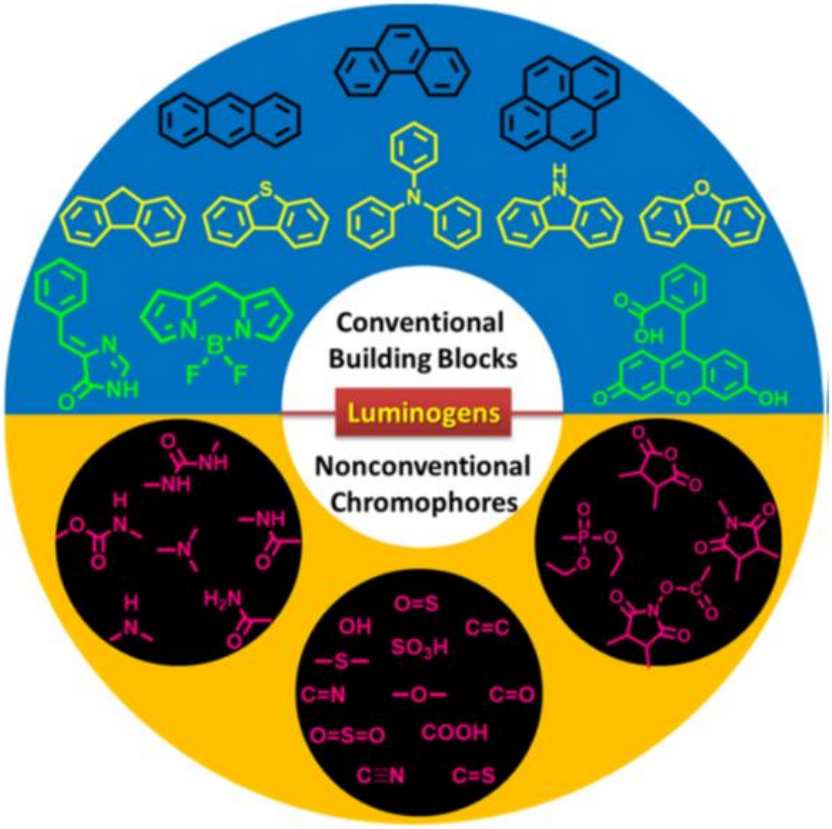


Figure 1.2: Schematic illustration of the building blocks of conventional and non-conventional chromophores. (Adapter from Zhang Yuan & Zhang, 2017)[39]

1.3. Objectives

Based on the novel intrinsic absorbance and luminescence properties from unconventional chromophores, there is a need to investigate more into the spectral origins in these systems and understand the role of the chromophores involved. However, our work is restricted to unconventional chromophores in proteins. Knowledge of these unconventional chromophores can help biophysicists look into the optical properties in proteins in an entirely new way, and these novel label-free intrinsic probes can be used not only in apoproteins but also in other proteins systems, be it in monomeric or oligomeric forms. Most of the work involved in my thesis mainly highlights understanding the optical phenomenon of these unconventional chromophores and their potential application in various biophysical phenomena. The major

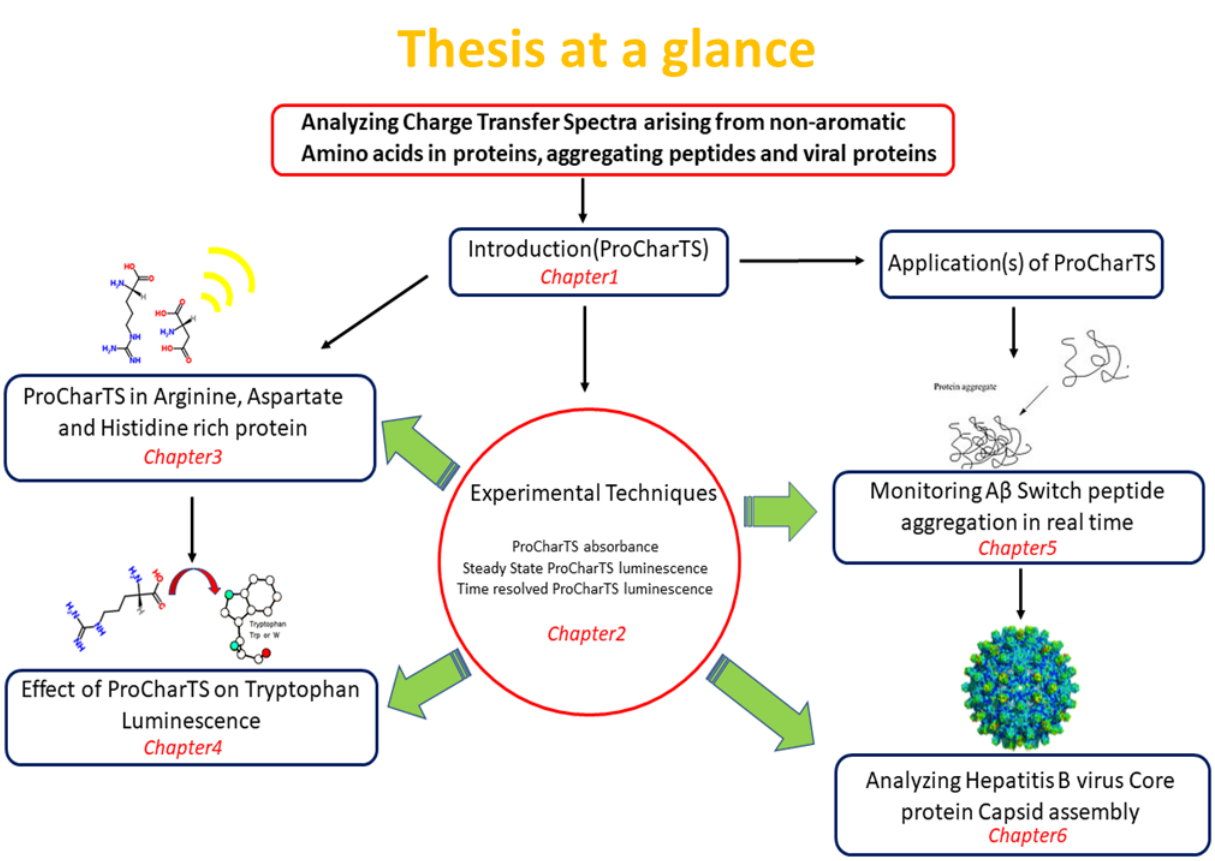
objectives of my PhD thesis are summarised below, and the discussion on each of these topics is carried out subsequently.

A: ProCharTS arising from Arginine, Aspartate amino acids observed in peptides & proteins which are devoid of Lysine residues.

B: Role of charged amino acids in sullying the fluorescence of Tryptophan in monomeric protein.

C: To monitor the early events of Amyloid beta (Aβ) Switch peptide aggregation in real-time using ProCharTS.

D. To monitor the Hepatitis B Virus Core protein (HBV Cp) capsid assembly in real-time using ProCharTS and extrinsic chromophores.



1.4. ProCharTS arising from Arginine, Aspartate amino acids observed in peptides & proteins which are devoid of Lysine residues

1.4.1. Protein Charge Transfer Spectra (ProCharTS) absorbance

As discussed in **Section 1.2.4**, the tremendous rise in the study of non-conventional chromophores from various biopolymers has gained importance recently. Here only the absorption phenomenon from apoproteins is being highlighted. Novel absorbance in the near UV-Visible region was reported not only in amino acids[27], L-Lysine-HCl[25] but also from small molecules such as HCL as well as NaOH[40]. Electron charge transfer in the hydrogen bond network in the water clusters as well as from the side chain of charged amino acids was thought to play an important role. Moreover, the origin of the novel absorbance features in the 300-800 nm region was also reported to arise from the amyloid fibrils[41]. Charge Transfer in the termini of these amyloid peptides was suggested to tune the optical properties characterizing a weak but broad absorption profile in the near-UV-Visible range, as shown in **Figure 1.3**.

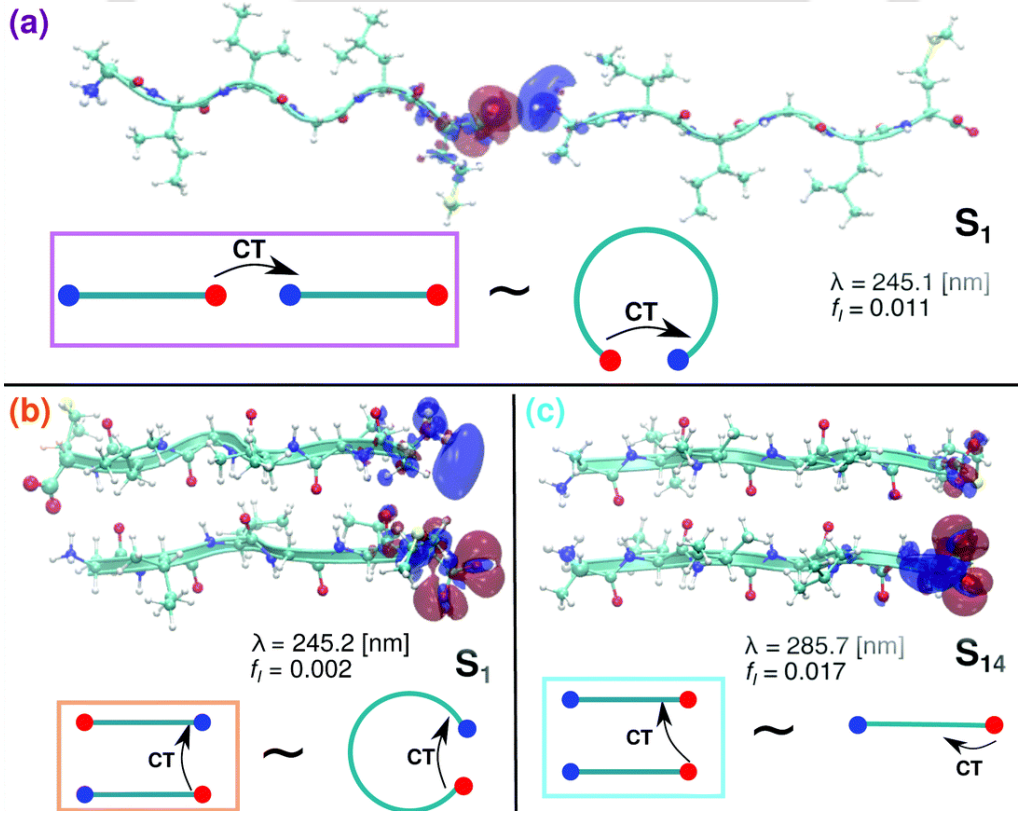


Figure 1.3: Iso-surface plot showing Charge Transfer between the ground and first excited states in the dimer of the segment of Aβ (AIIGLM) in linear parallel (a), antiparallel (b) and parallel (c) configurations. Electrons (red) and holes (blue) are highlighted. (Adapted from Jong et al., 2019)[41]

Similar absorbance features were also observed in larger proteins as well, devoid of any aromatic amino acids. Proteins with higher charged content were reported to show more absorbance in the UV-Visible region than in proteins with moderate charge content, shown in **Figure 1.4** *(left)*, hinting at the role of Charge Transfer in giving rise to the novel absorption feature[26]. Recently, the intrinsic UV-Visible absorption in a highly charged protein α_3C rich in Lysine (Lys) and Glutamate (Glu) with no aromatic amino acids was reported[23,42], as can be seen in **Figure 1.4** *(right).* Time-dependent density functional theory (TDDFT) calculations on MD simulation snapshots revealed various Lys-Glu, Lys-Lys and Glu-Glu contacts in the protein to be the reason for intrinsic absorption. This phenomenon has been postulated as ***Protein Charge Transfer Spectra (ProCharTS).***

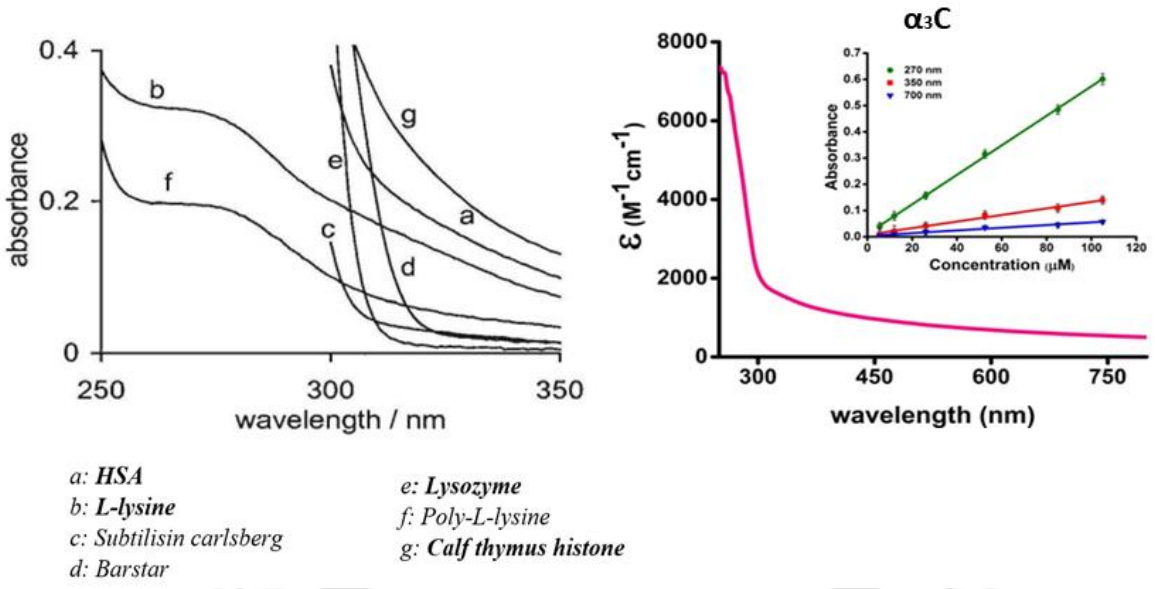


Figure 1.4: Novel ProCharTS spectra in charged-rich proteins (left) (Adapted from Homchaudhuri & Swaminathan, 2004)[26] and and α_3C (right) (Adapted from Prasad et al., 2017)[23]

ProCharTS absorption arises when excited state electrons in the side chains of cationic amine group in Lysine (acceptor) or anionic carboxylate head groups in Glutamate (donor) act as electron acceptors or donors for charge transfer, either to or from the polypeptide backbone or to each other in proteins as shown in **Figure 1.5.** The electronic transition from the HOMO (in donor) to the LUMO (in acceptor) occurs after electrons are photo excited, resulting in absorption spectra in the UV-Visible range (250–800 nm). Spatial proximity among the charged amino and carboxylate groups across the protein chain plays a major role and dictates the fate of the electronic transitions and optical features in ProCharTS.

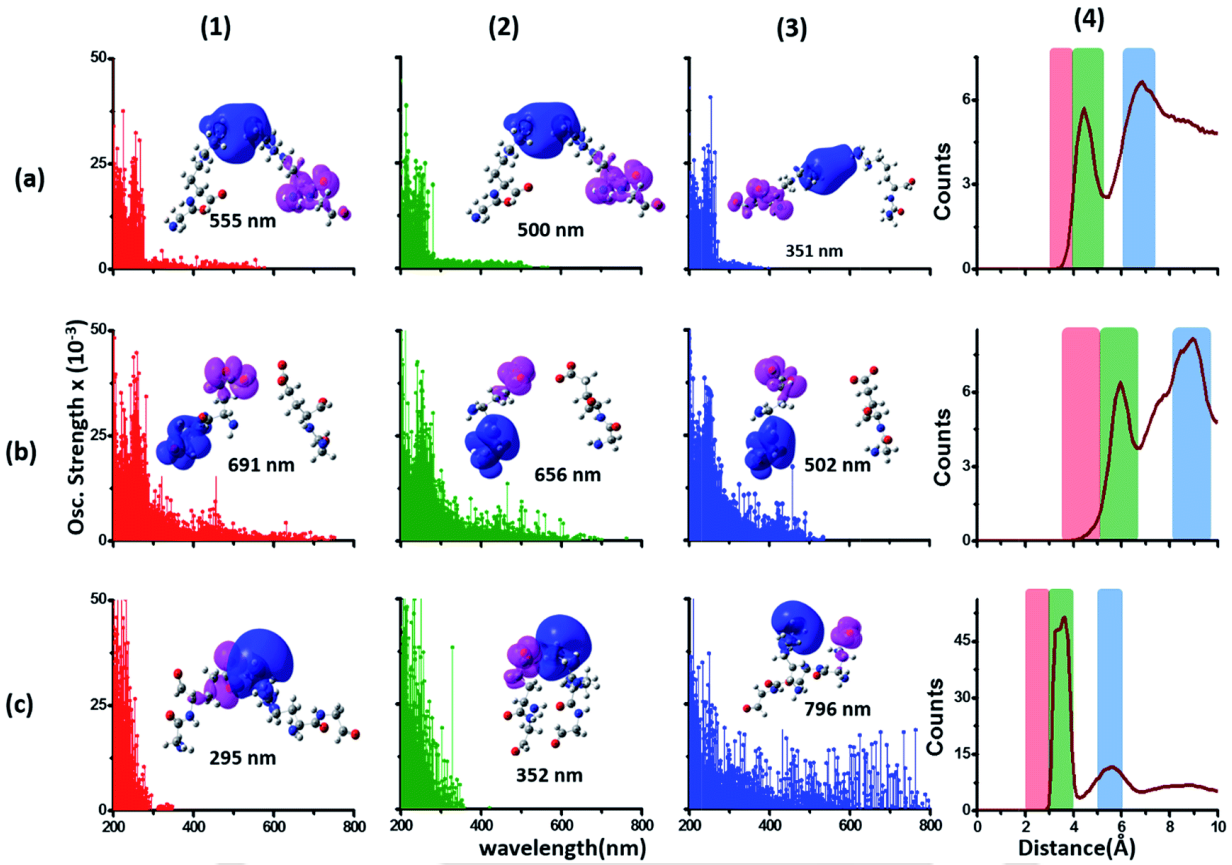


Figure 1.5: Simulated absorption spectra (oscillator strength vs wavelength) for dimers of a) Lys–Lys, b) Glu–Glu, and c) Lys–Glu. Subfigures 1, 2, and 3 represent the Charge Transfer at different distances; Subfigure 4 represents their respective Radial Distribution. (Adapted from Prasad et al., 2017)[23]

ProCharTS absorbance has been successfully utilized to analyse not only the conformational changes in proteins due to changes in pH or temperature but also was found to be a potential tool to monitor the aggregates of Human Egg White Lysozyme (HEWL)[43].

1.4.2. Protein Charge Transfer Spectra (ProCharTS) luminescence

Subsequent to absorption, the photochemical features of intrinsic novel luminescence in proteins were investigated by various groups. As already discussed in **Section 1.2.4**., the multiple reports of intrinsic deep-blue/blue luminescence in monomeric proteins, oligomeric proteins, amyloid aggregates, and even crystals of amino acids and proteins hint at multiple hypotheses. Similar to intrinsic absorbance, the origin of deep blue luminescence is quite elusive. The origin of such intrinsic fluorescence includes proton transfer in fibrils of aggregated protein via hydrogen bond[44] as shown in **Figure 1.6**, inter or intramolecular electron delocalization[33] shown in **Figure 1.7**, or from modification in the chemistry of amino

acids by oxidation[35]. Moreover, carbonyl double bonds present in proteins (HEWL, BSA, and β-lactoglobulin) were also thought to be the chromophore for such blue luminescence[45], suggesting the role of charge transfer involving carbonyl double bonds. Although the deep blue

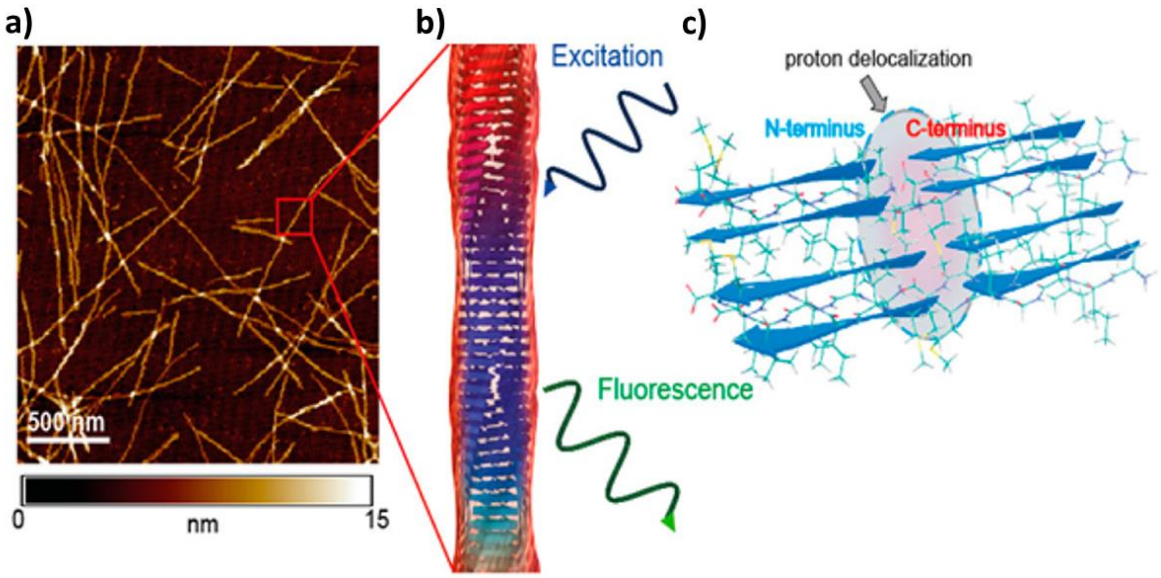


Figure 1.6: Structural arrangements of Aβ (AIIGLM). (a) AFM image (b) Schematic view of the 3D peptide arrangement in a protein fibril stacked by β-strands. (c) Cross-section of the magnified section from same two adjacent protofilaments. The shaded ellipsoid depicts predicted proton transfer. (Adapted from Pinotsi et al., 2016)[44]

luminescence is evident in various biological systems, mainly in clusters of charged species, amyloid fibrils or protein aggregates, and various phenomena such as Clustering Triggered Emission (CTE) or Aggregation Induced Emission (AIE) have been addressed, reports of similar deep blue luminescence have also been reported recently in various monomeric proteins

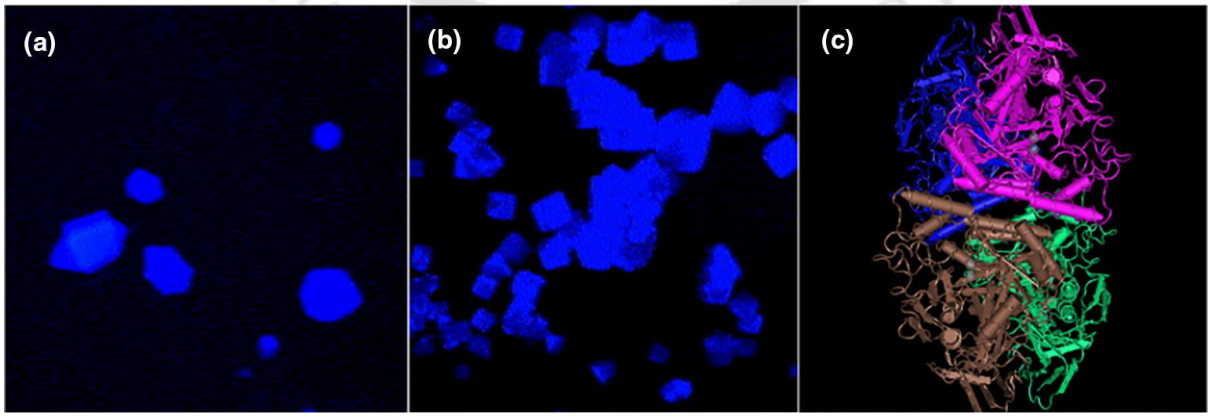


Figure 1.7: Images of (a) hydantoinase and (b) truncated (bab) Pyrococcus furiosus protein at 364 nm excitation. (Adapted from Shukla et al., 2004)[33] (c) Structure of D-hydantoinase. (Adapted from Xu et al., 2003)[46]

rich in charged amino acids[30], as shown in **Figure 1.8**. The monomeric nature of the protein was confirmed by the linear increase in luminescence with an increase in Human Serum Albumin (HuSA) concentration, and any possibilities of oligomeric or aggregated state to contribute to enhancement in luminescence was ruled out, contrary to earlier reports which argued only aggregation or oligomerization gives such blue luminescence[32,47]. Furthermore, recent work[40] involving concentrated aqueous HCl/NaOH solutions also displayed such intrinsic fluorescence, although no such aggregated fibrillary moieties were present. Moreover, the luminescence of other proteins in crystal form[33,48], other small molecules[27] lacking any aromatic amino acids and in a monomeric state has also been reported. Similar to ProCharTS absorbance, ProCharTS luminescence was proposed to be the origin of such novel luminescence from monomeric proteins and amino acids[30].

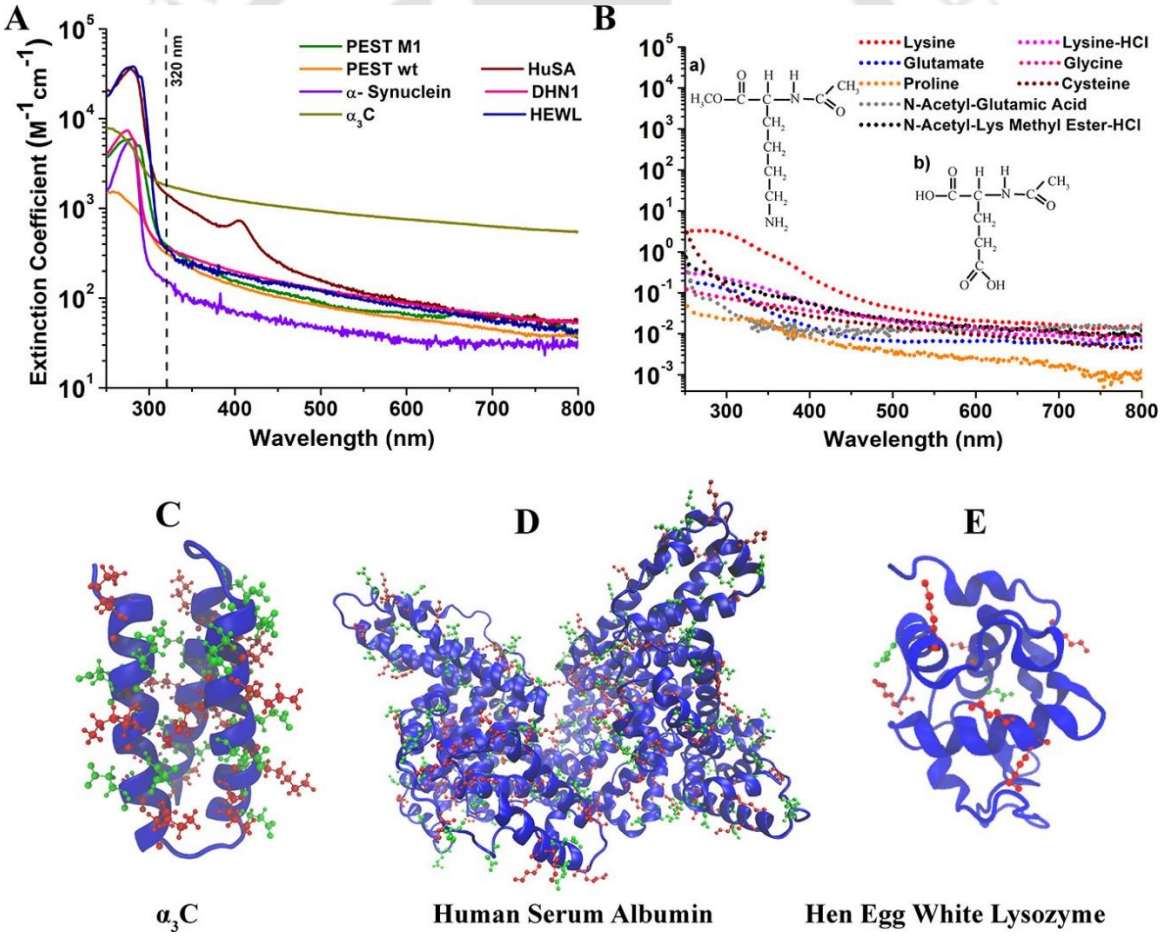


Figure 1.8: Extinction coefficient of (A) charged rich proteins and (B) charged amino acids. Subfigure C, D, and E highlight the charged content Lys (red), Glutamate(green) in α_3C, human serum albumin, and hen egg white lysozyme, respectively. (Adapted from Kumar et al., 2020)[30]

The concept of charge-transfer transitions between various charged species is thus raised to be the reason for blue luminescence. This novel luminescence is suggested to arise from the charge recombination phenomenon between excited state electrons and the created holes in the protein. Charge recombination is proposed to meditate by 1. direct electronic recombination to the hole or 2. delayed recombination with electrons hopping on the backbone of protein and then recombination either with the same or a new hole, as shown in **Figure 1.9.** ProCharTS in

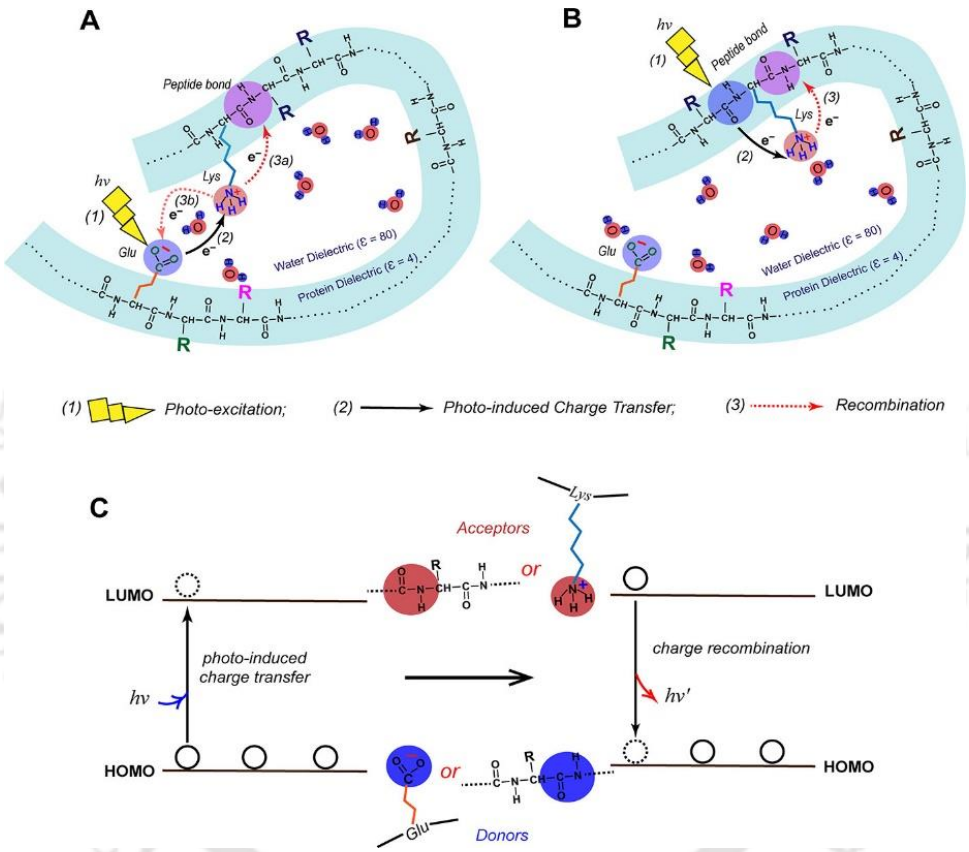


Figure 1.9: Subfigure A and B show the schematic illustrations of the two major pathways for photoinduced charge transfer and charge recombination. Subfigure (C) shows the schematic of charge transfer states and electronic transitions. (Adapted from Kumar et al., 2020)[30]

monomeric proteins is generally characterized by low quantum yield with multiple luminescence lifetimes. The origin of multiple luminescence lifetimes where charge recombination is mediated via multiple pathways also supports the charge recombination hypothesis. ProCharTS luminescence was found to be highly correlated with the product of molar extinction coefficient and quantum yield suggesting the major role of the molar extinction coefficient in enhancing intrinsic luminescence intensity given the poor luminescence quantum yield[30]. Thus, it was found that ProCharTS luminescence too is highly

dependent on the close spatial 3D proximity of the head groups of charged amino acids and is more pronounced in charge-rich proteins than in proteins with low or moderate charge content.

1.4.3. Evidence of ProCharTS in peptides and proteins devoid of Lysine

The majority of the monomeric proteins studied that showed intrinsic ProCharTS absorbance/luminescence are predominantly rich in lysine and glutamate amino acid residues. However, there are no reports of ProCharTS on proteins or peptides that lack Lysine or Glutamate while being rich in other charged amino acids like Aspartate, Histidine and Arginine. Although there are reports of novel intrinsic absorbance in the near UV-Visible region and also intrinsic blue luminescence from amino acid clusters of Arginine, Aspartate and Histidine in solution, there is no literature as of now that addresses and links the role of these charged amino acids in peptides or proteins. However, recently, TDDFT calculations on various amino acids have been reported and, as seen in **Figure 1.10,** ~~and~~ they have shown similar simulated optical features as that of Lysine and Glutamate in Arginine, Aspartate and Histidine amino acids[49]. It is hypothesized that ProCharTS exists as long as charged residues are involved, irrespective of the type of charge present in the protein.

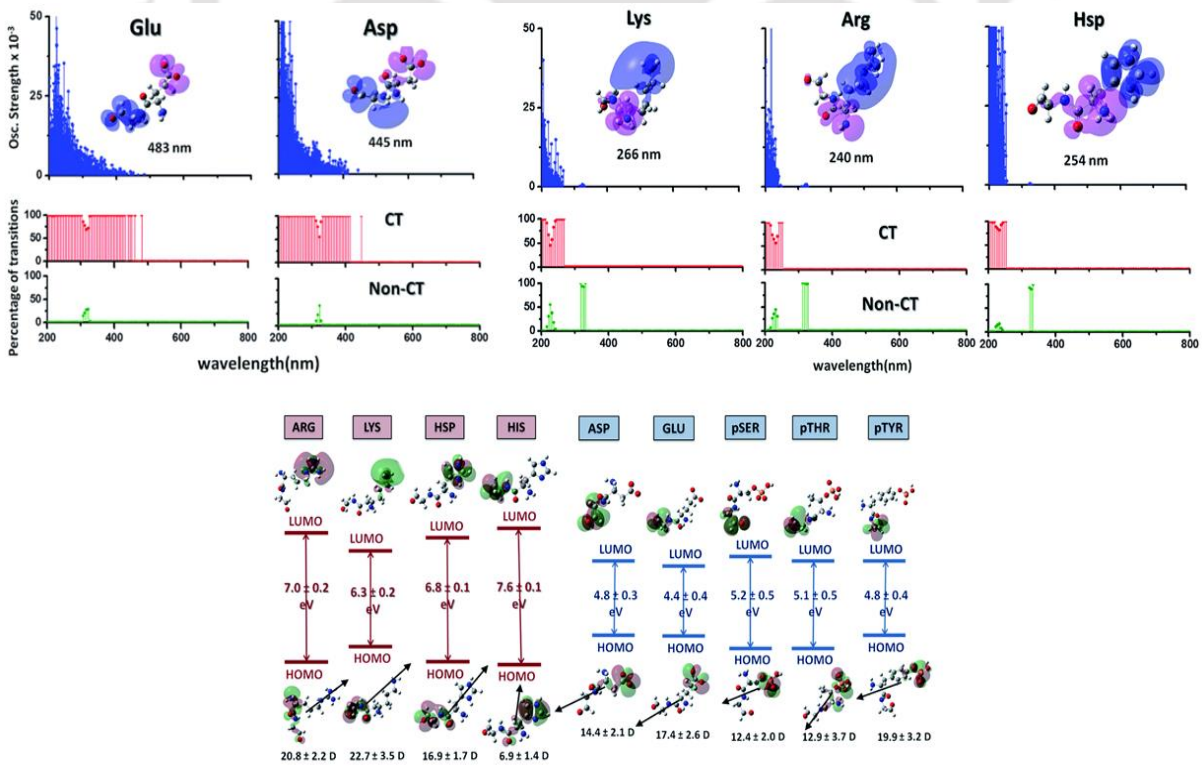


Figure 1.10: Simulated absorption spectra (oscillator strength vs wavelength) for various amino acids along with their average energy gaps in HOMO–LUMO and dipole moment values. (Adapted from I. Mandal et al., 2018)[49]

1.5. Role of charged amino acids in sullying the fluorescence of Tryptophan in monomeric protein

1.5.1. Introduction

As already mentioned in **Section 1.2**, spectroscopic techniques like electronic absorption or fluorescence are often the most common methods to gain insights into the structure and dynamics of various conformations of protein and study their structure-function relationship. Although extrinsic (fluorescent dyes) are frequently used to yield information about protein dynamics, like in FRET, some inherent problem with extrinsic probes lies in the fact that they can sometimes alter the protein structure and conformation as they are conjugated to the protein surface[50,51]. Intrinsic chromophores in proteins such as aromatic amino acids, prosthetic groups, and peptide bonds, however, are non-invasive and hold an advantage over extrinsic dyes in this regard. Among the aromatic amino acids present in protein, Tryptophan is the most extensively used probe to monitor the protein dynamics due to its relative higher absorbance and fluorescence features as compared to Tyr and Phe. The indole group present in Trp is thought to be the chromophore responsible for the emissive properties[2]. Moreover, Trp is highly sensitive to solvent and local environmental changes as well as local fluctuations and provides precise information about the surrounding of the chromophore. The emission spectrum of Trp shows varied maximum emission (λ_{max}) from about 310−350 nm, multiple fluorescence lifetimes components ranging from 50 ps to 8 ns, and relatively high fluorescence quantum yields (Φ) with a maximum of around 0.35. Proteins which contain buried Trp inside the hydrophobic core when excited at 295 nm wavelength show blue-shifted fluorescence maxima, whereas marginal exposure of Trp in protein to its surface displays red-shifted emission maxima due to solvent exposure. Moreover, FRET and quenching studies of Trp also yield valuable information regarding the overall conformation of proteins as well as its interaction with other biomolecules.

1.5.2. Factors affecting indole fluorescence

Single Trp residue in a protein typically gives rise to fluorescence decays that are generally fitted by sums of 2 to 4 exponential terms, i.e., fluorescence lifetime components with varying respective amplitudes. N-Acetyl-tryptophan-amide (NATA), which is a tryptophan analogue however, displays only one fluorescence lifetime component in water. The reason why

Tryptophan displays multiple fluorescence lifetime components in proteins is quite controversial. The opinions on this phenomenon can be broadly divided into two categories. First and the most generalized explanation is the "classical rotamer model" that emphasizes ground-state heterogeneity, which suggests that discrete subpopulations of different Trp χ_1 and χ_2 rotamers, as shown in **Figure 1.11** exhibit different decay times that arise from the various interactions of the different Trp ground state conformers within the protein matrix[52–54]. Similar observations were also made regarding the different rotameric structures in the *Cα-Cβ* bonds[55,56]. The second type of explanation is in terms of an excited-state process instead of the more popular ground-state heterogeneity theory[2,57]. In this case, it is assumed the excited chromophore can reorient from its initial configuration to a whole new fluorescent species after electronic excitation. Although both these popular models find success in model studies, they are not justifiable in many circumstances. Certain other models thus were also proposed for Trp multi-exponentiality. In these cases, reversible generation of a nonfluorescent species that return to the emissive state is thought to be involved[58]. Here, Electron transfer from the proximal excited tryptophan indole group to a nearby quenching residue results in a subsequent charge recombination phenomenon, thereby displaying multiple lifetime components.

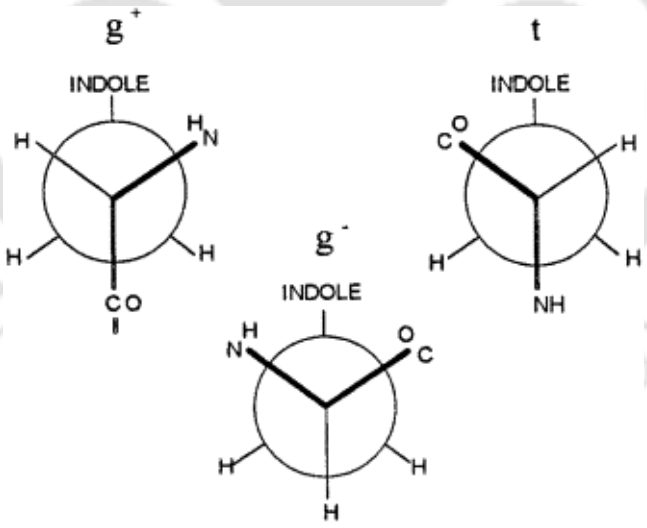


Figure 1.11: Rotamers of the indole side chain present in Tryptophan. The rotation of the indole group along the C_α-C_β bonds is indicated by the Newman projections. (Adapted from Clayton & Sawyer, 1999)[53]

Various reports on the effect of Electron Transfer have been observed on Trp fluorescence. Reports of efficient electron transfer from the peptide group present in the peptide bond near to indole were also reported[59,60]. Some groups also observed quenching of Trp fluorescence from Histidine cation and disulphide bonds[61,62]. Moreover, the role of proximal charged amino

acids Lys, Glu, Arg, and Asp in the presence of the indole group was also thought to be a plausible reason for Trp quenching by excited state proton transfer[63–65].The charged groups in the vicinity of the indole group in proteins were shown to shift the emission maxima of Trp through long-range coulomb interactions[66].The proximity and orientation of these amino acid side chains surrounding the indole ring highly affected the Trp fluorescence.

1.5.3. Effect of ProCharTS on Tryptophan fluorescence

Recently, it has been reported that ProCharTS may play a role in sullying the fluorescence from Tryptophan, where Charge Transfer between the indole ring in Trp and the excited electronic states of various charged charge amino acids are expected[30]. It can thus create a heterogenous system where recombination phenomena could take place. The work highlighted the role of the interference from ProCharTS luminescence in the collective influence of a group of charged amino acids in amino acid clusters as well as in protein on the Trp fluorescence. Spectral overlap of both excitation and emission from Trp with ProCharTS was reported, as can be seen in **Figure 1.12**. Since both ProCharTS absorbance and luminescence pattern tend to overlap in

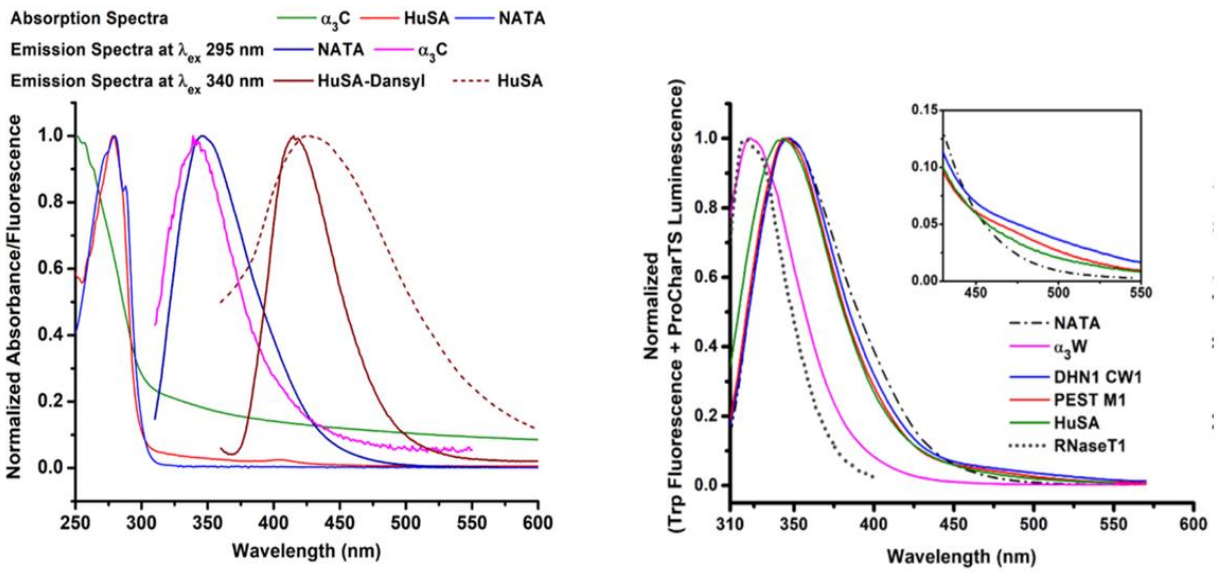


Figure 1.12: Spectral overlap of indole fluorescence with ProCharTS absorbance and luminescence (left). Boarding of Trp fluorescence spectra in the presence of ProCharTS (right). (Adapted from Kumar et al., 2022)[67]

the same wavelength range of Tryptophan, the contribution of the additional optical properties from ProCharTS to the Tryptophan fluorescence in charged-rich proteins was suggested. Moreover, a characteristic red-shifted spectrum with broadened emission features in a

relatively longer wavelength region was observed. Free NATA in solution in the presence of charged proteins displayed similar multiple fluorescence lifetimes as that of Tryptophan but less so in moderately rich charged proteins, as shown in **Figure 1.13**. Also, proteins, in comparison to free amino acids, sully the intrinsic NATA fluorescence manifold as the free charged amino acids cannot create a realistic protein-like environment where charged clusters

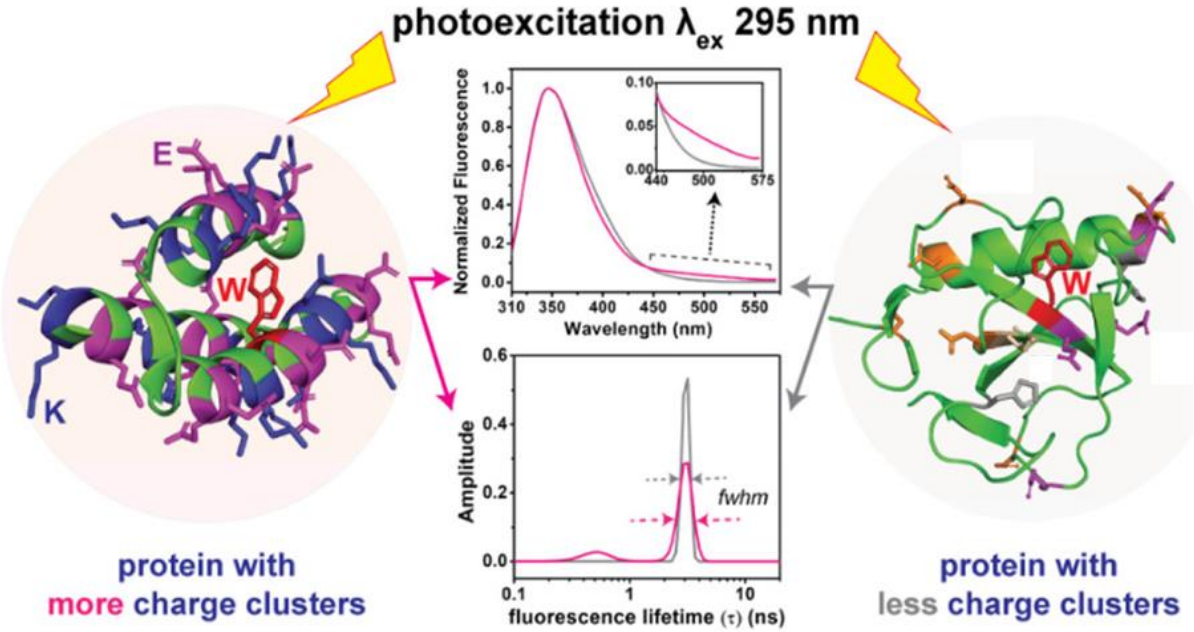


Figure 1.13: Schematic illustration of the effect of ProCharTS on the Tryptophan fluorescence in protein with high charge content. (Adapted from Kumar et al., 2022)[67]

in proteins are in close proximity with the indole group. The effect of ProCharTS on Tryptophan was observed mainly in Lysine-rich proteins as compared to proteins rich in other charged non-aromatic amino acids. One reason that was argued was that Lysine, among the charged non-aromatic amino acids, showed high molar absorptivity in the UV-Vis spectrum with superior luminescence properties. Moreover, the effect of charged amino acids was not only restricted to intrinsic chromophores like Tryptophan but also on extrinsic dyes like Dansyl Chloride when labelled in charged proteins. A similar influence on Dansyl fluorescence was observed with quenched fluorescence, broader emission and the addition of a shorter fluorescence lifetime component to the relatively higher fluorescence lifetime of Dansyl. Therefore, in this study, we tried to further address the effect of ProCharTS from other charged amino acids Arginine, Aspartate and Histidine on Trp. Moreover, proteins and peptides containing exclusively these residues were considered in this study which will be discussed in later sections.

1.6. To monitor the early events of Amyloid beta (Aβ) Switch peptide aggregation in real-time using ProCharTS

1.6.1. Introduction

Alzheimer's disease (AD) is a severe neurodegenerative disorder where insoluble amyloid fibrils deposit in the neuropil as ordered aggregates and is caused by Aβ40 or Aβ42 peptides. As per the 2015 Alzheimer world report, AD has adverse effects on many individuals, close to 46 million people and has cost more than 800 billion dollars globally[68]. Amyloid precursor protein (APP) cleavage forms the Aβ42/Aβ40 peptides, which are the main contributors to AD. APP gets cleaved in the fibrillogenic C terminal domain by the γ secretase within the membrane region. This γ secretase cleavage is not specific and may cause many isoforms where the more common fragments being Aβ42 and Aβ40. The less amyloidogenic N terminus gets cleaved by the extra cellular β secretase.

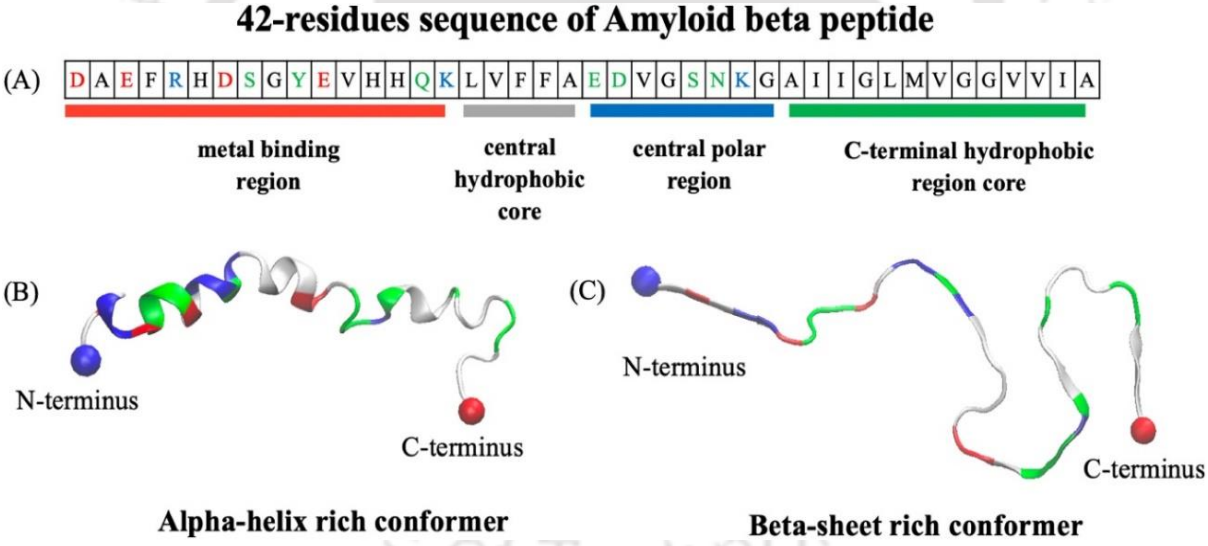


Figure 1.14: Subfigure A shows the sequence of Aβ42; subfigures B and C show the NMR-resolved structure of Aβ42 α-helix (1Z0Q) and β-sheet (2NAO) rich conformers, respectively. (Adapted from K. Y. Wu et al., 2021)[69]

The composition of the monomeric Aβ42/Aβ40 peptides possesses various different types of secondary structural varieties. The fragment $A\beta_{17-21}$ constitutes the hydrophobic core. While both fragments $A\beta_{28-42}$ and $A\beta_{9-21}$ are able to show alpha helix/beta-sheet secondary structure, the $A\beta_{28-42}$ fragment is more prone to form the beta sheets. More specific information regarding small fragments is predicted from structural studies and have found that the fragments $A\beta_{33-42}$ and $A\beta_{17-21}$ show the greatest hydrophobicity while the C-terminus $A\beta_{33-42}$ showed a much

higher probability of forming β-sheet shown in **Figure 1.14.** Two β-turns are predicted to constitute the full-length Aβ and are found between $A\beta_{6-8}$ and $A\beta_{23-27}$ fibrils[70,71].

1.6.2. Biological relevance of the oligomeric stages of the peptide Aβ

The initial fibrillary oligomers play a detrimental role in the pathogenicity in AD and need to be researched further[72,73]. For example, when cerebral microinjections of Aβ monomers and oligomers were injected in rats *in vivo*, it affected long-term hippocampal potentiation, whereas no such effects could be seen with microinjections containing mature amyloid fibrils[74]. Similarly, recent evidence also points out the very fact in the case of humans too, the small and soluble Aβ oligomers in cerebrospinal fluid (CSF) of AD patients are highly correlated with AD onset than in mature insoluble fibrils. Moreover, Aβ oligomers play a vital role in the death of neuronal cells and brain disorders affecting the ionic flow, loss of synaptic function and kinase activity[75,76]. Thus, understanding Aβ aggregation and inhibiting it in its early stages by identifying potential drugs is of utmost importance for treating AD and other related amyloidogenic diseases.

1.6.3. Reason for the choice of the $A\beta_{16-22}$ (KLVFFAE) fragment of the peptide

Different amyloid fibres possess a common beta structure, and it has been hypothesized that steric zippers, a motif of tightly stacked beta sheets, not only serve as the spine of the protofilaments but also can mediate the interactions between protofilaments that associate to form mature fibrils[77]. Multiple segments within $A\beta_{11-25}$ and $A\beta_{27-42}$ are predicted either to self-associate into homotypic or to form heterotypic steric zippers. These predicted heterotypic interactions correlated well with ss-NMR studies of Aβ40[78,79]. Thus, understanding the nature and interaction of such steric zippers can give insights into the actual mechanism of full-length Aβ peptide aggregation. Another reason to choose the segment $A\beta_{16-22}$(KLVFFAE) is that it forms an anti-parallel beta-sheet structure similar to that in the full-length Aβ. Such anti-parallel conformation is also quite commonly associated with the "Iowa" Aβ mutant, a familial early-onset AD[80,81]. Thus, $A\beta_{16-22}$ being one of the shortest sequences to form aggregates and is easier to synthesize and form stable fibrils, whereas other fragments like $A\beta_{10-23}$ and $A\beta_{29-42}$ can form very short fibrils, which are quite insoluble or relatively longer fibrils. Thus, $A\beta_{16-22}$ turns out to be an ideal and alternate candidate for aggregation study as special and costly synthetic and purification skills are required to prepare the full-length Aβ.

1.6.4. Aβ switch peptides

One potential problem with studying Aβ peptide is that it is extremely amyloidogenic in nature, thus making it hard to synthesize and purify extremely challenging, and in-depth information regarding its mechanism of early events of aggregation pathway is still limited[82]. Therefore, developing an alternate model is of particular interest, which is reliable, can be readily accessible and can facilitate the study of initial amyloidogenic events to look for potential Aβ inhibitors. During the last decade, reports emerged[83,84] of a precise mechanism to study early events of amyloid aggregation by introducing molecular switches into polypeptides which are termed Switch peptides. Here, the switch peptide (SwP) of $A\beta_{16-22}$ (KLVFFAE) is being studied here in this study. Its principle is simple and based on a host-guest system where two β-sheet-promoting $[Leu\text{-}Ser]_n$ sequences (host) flank our selected Aβ fragment comprising $A\beta_{16-22}$ (KLVFFAE) (guest) where the main role of the $[Leu\text{-}Ser]_n$ is to help forming oligomers thus promoting fibrillization, which is illustrated in **Figure 1.15**. Switch elements were placed strategically into the guest sequence, and this was done by placing them outside of the target sequence in the peptide so that the sequence of our guest was not disturbed. When pH is changed from acidic pH (~pH 2) to neutral or slightly basic pH (~pH 6.8-7.2), the O to N acyl

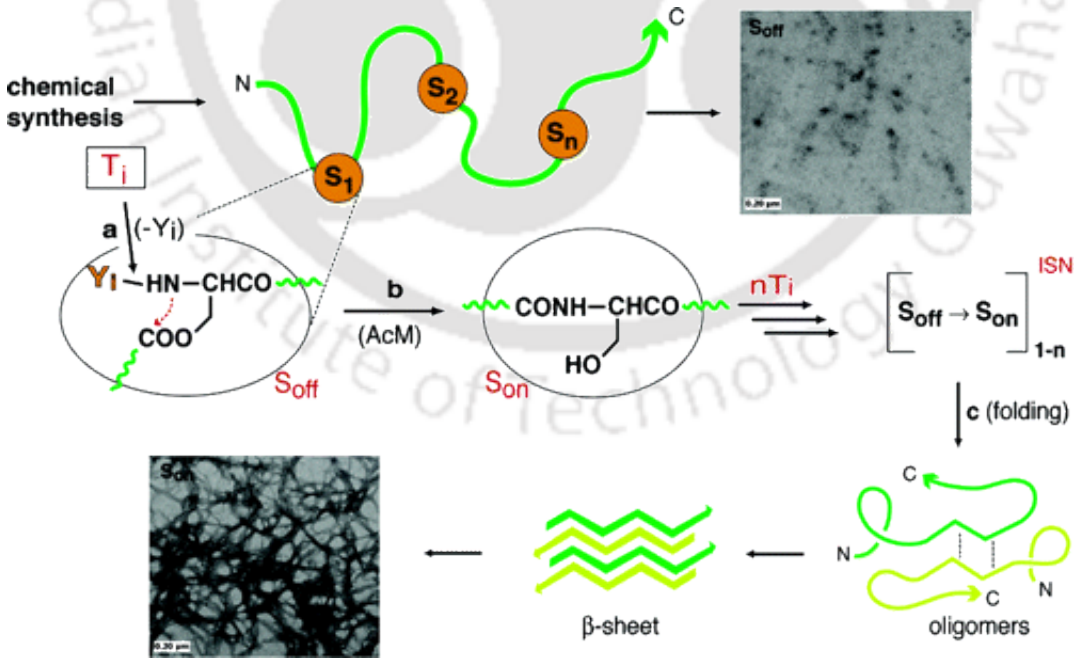


Figure 1.15: Proposed mechanism of Switch peptide aggregation. (Adapted from Dos Santos et al., 2005)[84]

migration in the molecular switch propels the initially random coiled peptides to convert into β-sheet and trigger self-assembly. In other words, the speed of the acyl migration by restoring

the peptide backbone from its respective molecular switch can be tuned very precisely by adjusting from acidic pH (~pH 2) to neutral or slightly basic pH (~pH 6.8-7.2). Thus, this efficient host-guest switch peptides model specifically serves our interest in systematically characterising the early folding mechanism of Aβ peptides aggregation.

1.6.5. Conventional techniques to study Aβ aggregation

Many techniques like solution NMR, novel methods based on nanopore technology[85], and fluorescence techniques with extrinsic fluorophores such as ThT, ANS[51,86,87] are used to analyse the kinetics of aggregation. Also, some other techniques[78,88–92] which are of routine use to perform structural studies of soluble Aβ are analytical ultracentrifugation (AUC), Dynamic Light Scattering, Circular Dichroism (CD) and Fourier-transform infrared spectroscopy (FTIR), Atomic force microscopy (AFM) techniques, Transmission Electron Microscopy (TEM), ss-NMR and FTIR. Moreover, antibody-based techniques, and aptamers-based techniques[93] have been used recently to monitor the aggregation of Aβ peptides. But these methods have some limitations as some are quite expensive, labour intensive and not feasible for easy detection, while some do not provide information on molecular level interactions and only give information on the gross shape and size. Moreover, extrinsic fluorophores such as Thioflavin T (ThT) binding[50,51] and similar other dyes used for Aβ aggregation are invasive in nature and can interact with the aggregation process and alter the kinetics of the Aβ aggregation.

1.6.6. ProCharTS as a novel approach to study Aβ aggregation

To overcome the limitations from the conventional techniques, there is a need for an alternative label-free approach using intrinsic probes to monitor Aβ peptides aggregation. Recent reports of aggregation of HEWL showing increased ProCharTS absorption upon aggregation[43] and intrinsic fluorescence or intrinsic deep-blue/blue luminescence arising from oligomeric proteins[32] have been encouraging in this aspect, as shown in **Figure 1.16**. Thus, it would be exciting to see the role of ProCharTS in peptide aggregates as no such literature involving ProCharTS monitoring Aβ peptides is present till now. Light scattering from aggregation is not expected to interfere with ProCharTS since only the soluble oligomeric aggregates in the initial stages of aggregation are involved in our study, and scattering is often associated with higher order insoluble mature fibrillar aggregates.

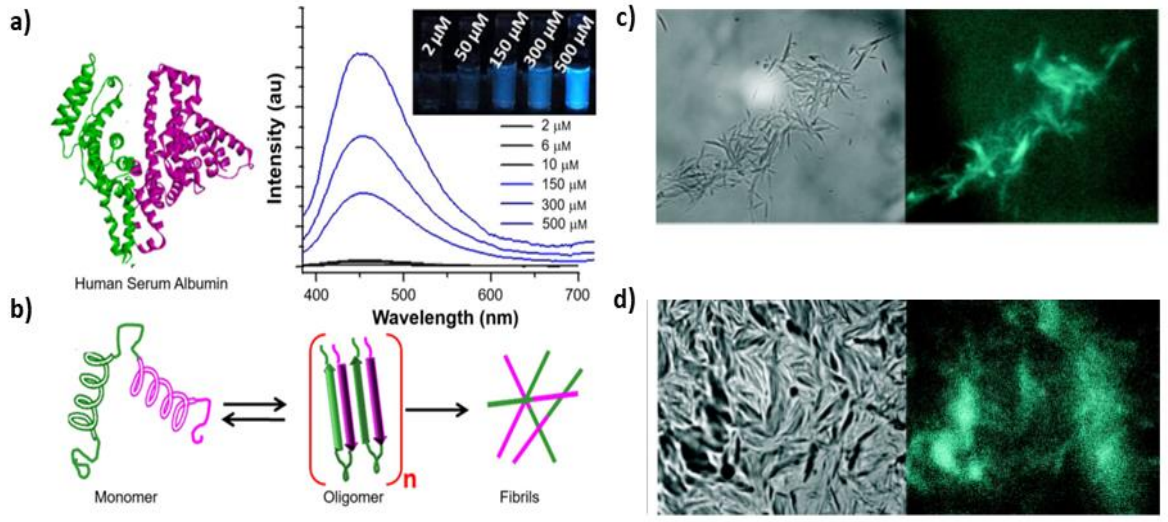


Figure 1.16: The structure of HSA along with its emission spectra at 2 to 500 μM concentration upon excitation of 375 nm (a) and the schematic illustration of fibril formation (b). (Adapted from Bhattacharya et al., 2017)[32], Subfigures c) and d) show fluorescent microscopy images of amyloid fibrils in completely hydrated and dehydrated states, respectively. (Adapted from Sharpe et al., 2011)[94]

An aggregation is basically an intermolecular event where proteins/peptide assemble to form oligomers and finally forms fibrillary structures. Since during $A\beta_{16\text{-}22}$ aggregation, Phe residues

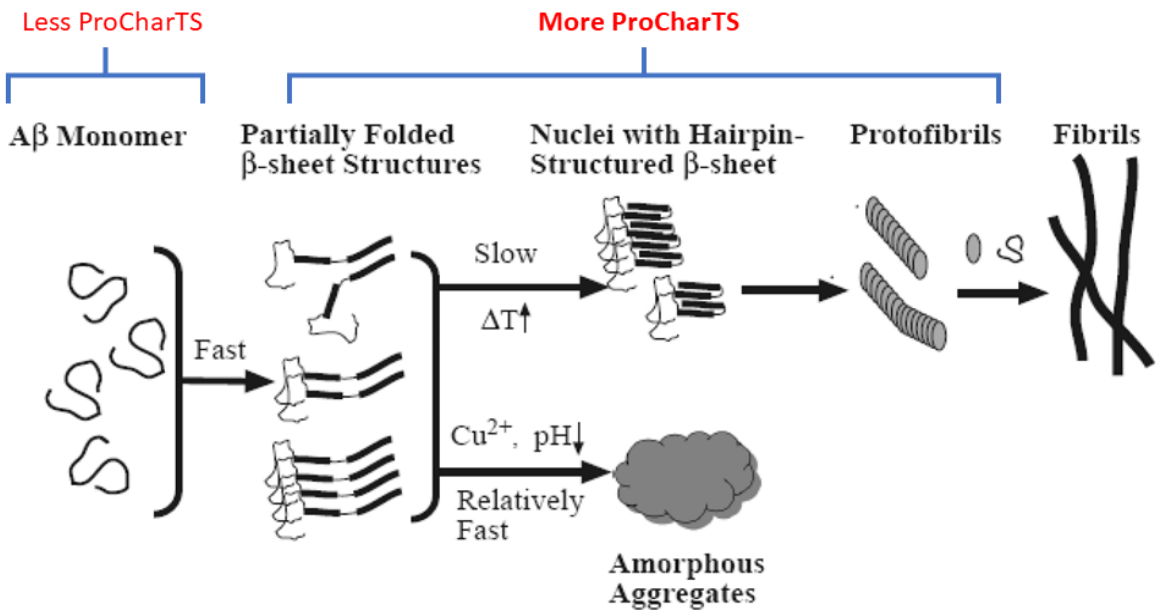


Figure 1.17: Illustration of possible ProCharTS in the oligomeric states of Aβ. (Adapted from Jiang et al., 2012)[95]

are stacked together by forming the hydrophobic core, the charged residues are expected to come at close proximity, and ProCharTS is expected to arise from it, which is expected to be absent in its monomeric state as illustrated in **Figure 1.17**. Interestingly, it has been found by

multiples groups[91,96] that the arrangement of each individual anti-parallel β-strand in any β-sheet has a hydrogen bonding distance of 4.8 Å, and each sheet has a spacing of 10-11 Å. Since these distances of 4.8 Å and 10-11 Å (each individual anti-parallel β-strand and between β-sheets, respectively) are in close 3D proximity, we can expect Lys/Lys or Lys/Glu interactions in their flanking ends of $A\beta_{16-22}$. So, it will be worth determining whether ProCharTS absorbance and luminescence can detect the time-dependent growth of aggregates in Aβ peptides.

1.7. To monitor the Hepatitis B Virus Core protein (HBV Cp) capsid assembly in real-time using ProCharTS and extrinsic chromophores

1.7.1. Introduction

Hepatitis B virus (HBV) is a small enveloped parsimonious virus that generally infects the hepatocytes and belongs to the hepadnavirus genus. It is transmitted from the parents to the child, through sexual means or through any other parenteral pathway. HBV virions are 40-42 nm in diameter, having double layered morphology. The outer lipoprotein layer contains three distinct glycoproteins called surface antigens[97,98]. Inside the outer envelope is the viral core[99] which contains a relaxed-circular, partially double-stranded DNA with a genome size of 3.2 kb and a polymerase, which acts as reverse transcriptase in synthesizing the viral DNA in virus-infected cells[100]. The partially double-stranded DNA consists of a full-length (-) strand and an incomplete (+) strand which replicates via an intermediate single-stranded pregenomic RNA (pgRNA). When an individual is infected, the DNA genome converts itself into a covalently closed circular DNA (cccDNA) in the nucleus and mRNAs are transcribed from it[101,102]. The DNA has four open reading frames[97,99]. First is the presurface/surface (preS1/preS2/S) region, and it encodes the S protein (HBsAg), which constitutes the bulk of the envelope proteins: the M (preS2) protein and the L (preS1) protein which help the virus to bind to host-cell receptors The PC/CP regions encode hepatitis B core antigen (HBcAg) or core protein (HBV Cp), which encapsids the viral RNA and hepatitis B e antigen (HBeAg) which helps in modulation of the host immune response. The third coding region encodes the viral polymerase which helps in DNA synthesis as well as RNA encapsidation[103]. The forth reading frame encodes the viral X protein (HBx) helps in the modulation of host-cell signal transduction[104].

1.7.2. Biological relevance of HBV Core protein (HBV Cp)

The capsid assembly in the HBV virus is driven by the HBV Cp, which encapsulates the viral RNA inside these empty capsids and aids in synthesising relaxed circular DNA (rc DNA) from pgRNA as well as for virus maturation and release[105–107]. The root of chronic HBV infection is the covalently closed circular DNA (cccDNA). Recently it was reported through Chromatin Immuno-Precipitation (ChIP) assay that HBV Cp does interact with the CpG(cytosine-phosphate-guanine) islands of ccc DNA. HBV Cp thus indirectly helps in recycling the mature rc DNA to the nucleus, thus replenishing the cccDNA reservoir to maintain chronic infection levels[108]. Since treatment with nucleoside analogues (NAs) which prevent new virions formation only sporadically, leads to loss of viral infection[109], there is a need for other novel approaches to cure chronic HBV. Recently, Core Protein Allosteric Modulators (CpAMs) that drive HBV Cp assembly[110] showed depletion of cccDNA-bound HBV Cp and decreased levels of other viral proteins[111]. Also, the HBV Cp capsid is highly immunogenic and is reported to induce both B and T cell responses[112]. Thus, understanding the HBV Cp protein and its assembly can be of immense help in eliminating chronic HBV infection. Moreover, human HBV is found to be highly specific to organs and species which challenges its analysis and therapy under *in vivo* conditions[113]. So, for the initial screening of drugs and understanding their mechanism, simpler models like *in vitro* studies are generally carried out. However, human HBV is not known to infect cell lines *in vitro*[113], whereas the HBV Cp assembly process, which is an important step in HBV infection, can be studied quite easily even with or without the involvement of nucleic acids. It has been reported that although most virion is uniform, some may appear small or outwardly aberrant[114]. A recent study too illustrated that as much as 90% of capsids might be empty[115]. Thus, targeting the empty and filled capsid assembly formation in HBV infection provides a very efficient yet simplistic model for its cure and provides better alternative therapies to cure HBV infection.

1.7.3. HBV Core protein (HBV Cp) structure and capsid assembly

The HBV Cp is a 22 kDa protein exists as a homodimer in the cytosol and consists of 183 or 185 amino acid (aa) residues depending on the genotype. It is one of the proteins in the HBV genome which is highly conserved among all the HBV isolates[116]. HBV Cp is basically comprised of 5 alpha helices. Helices α_3 and α_4 comprise the dimerization interface of each monomer which pack together to form a stable dimer by a disulphide bond, whereas the other

three helices α_1, α_2, and α_5 form the base of the protein, giving the template for the hydrophobic dimer-dimer contacts during HBV capsid assembly[117]. From the functional aspect, HBV Cp can be divided into two functional domains. The first 149 residues, Cp149, alone can assemble spontaneously into capsids which are morphologically similar to capsids formed from full-length HBV, as shown in **Figure 1.18**. The C-terminal domain Cp150-183 (34 residues) is highly charged, rich in arginine residues that help in RNA packaging as well as in reverse transcription of the pgRNA-Pol complex[118].

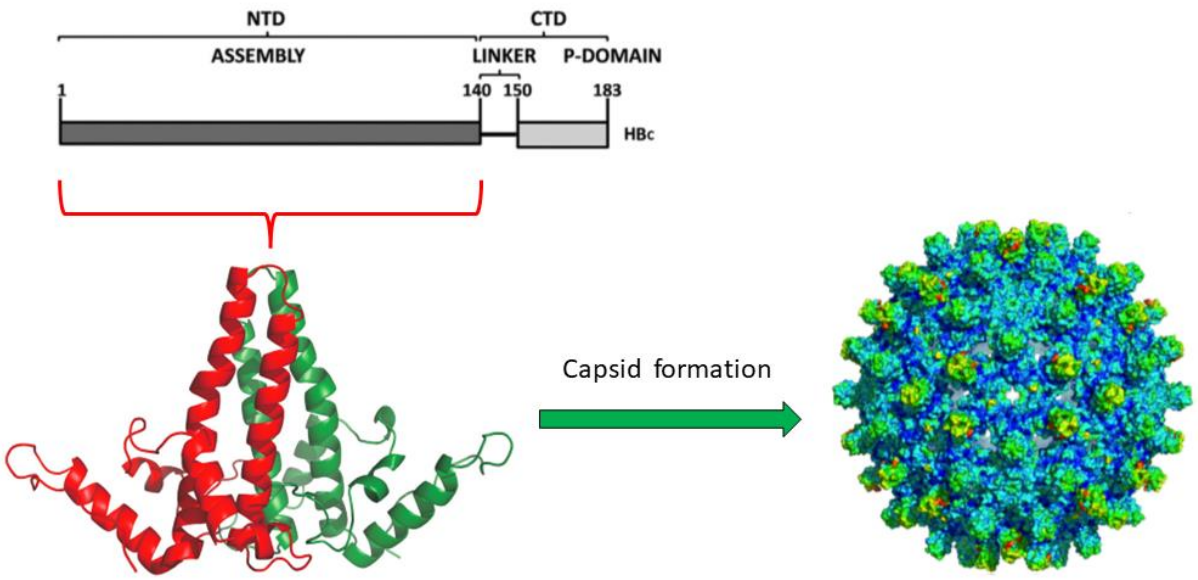


Figure 1.18: Schematic illustration of structure of HBV Cp 149 and its assembled capsids.

The HBV Cp capsid is formed in the cytosol and is initiated by the homodimerization of two HBV Cp, where a disulfide bond between Cys-61 residue stabilizes the protein[119,120]. The dimer then continues to oligomerise in an icosahedral symmetry forming capsids: (i) of 90 dimers (T = 3) with a diameter of 30 nm and (ii) a larger capsid of 120 dimers (T = 4) of about 34 nm in diameter. However, the T = 4 capsids are found more commonly in virally infected individuals, although both capsids are present simultaneously[121]. Moreover, the capsids can be either empty or filled with nucleic acids. *In vitro* assembly of Cp183, regardless of the absence or presence of heterologous RNA, showed morphologically indistinguishable capsids[122]. A high level of salt favours capsid assembly by disrupting electrostatic repulsions and forming weak interdimer interactions that stabilize the capsids[117,123]. In fact, such dimer contacts are characterized by the burial of ~1,700 $Å^2$ surface area during capsid formation. The assembly

kinetics is a 2-step phenomenon where 'n' subunits go through a rate-limiting nucleation step followed by rapid addition dimers each at a time.

$$120\ \text{Cp149} \rightarrow \text{nucleus (n)} + (120-\text{n})\ \text{Cp149} \rightarrow \text{capsid.}$$

As revealed by cryo-electron microscopy (cryo-EM), the capsid morphology indicates that the capsids have pores about 12–15 Å in diameter through their surface[117]. The entry and exit of nucleotides during DNA synthesis and degradation of the pgRNA are mainly facilitated by these pores.

1.7.4. Conventional tools to detect HBV Cp capsid assembly

Here, we highlight recent advances in virus capsid characterization and review both ensemble and single-particle methods. Virus capsid characterization and assembly have been studied *in vitro* by 90° light scattering[124], Transmission Electron microscopy (TEM)[122], X-ray crystallography[117], Dynamic Light Scattering (DLS)[125] and have contributed to our understanding of HBV virus capsids formation. Recently, Mass spectroscopy is found to provide effective temporal resolution to analyse HBV virus assembly in real-time[126]. Charge Detection Mass Spectroscopy (CDMS)[127] has emerged as a powerful tool to characterize the intermediate products of the HBV assembly. Also, resistive-pulse sensing, based on fluid-based particle counters, has been developed. As a matter of fact, currently, the highest resolution obtained for HBV capsids is performed with a combination of multi-cycle resistive-pulse sensing and with CDMS[128]. Atomic Force microscopy (AFM) has also been used to analyze HBV morphology[129]. Moreover, cryo-EM techniques have been used recently to analyse the 3D structure of HBV[130]. This eliminates the inherent problems faced due to dehydration and adsorption of the sample damaging viral capsids in conventional microscopy. In cryo-EM, samples are not fixed or stained but are frozen-hydrated, thus maintaining the 3D structure of the viral capsid. Separation technique such as Size Exclusion Chromatography[124] has also been used extensively to characterize HBV dimers from capsids. Similarly, Multiangle light scattering (MALS)[131], small-angle X-ray scattering (SAXS)[132], Nano particle Tracking analysis (NTA)[133], solid-state NMR[134], fluorescence correlation spectroscopy (FCS)[135], differential scanning calorimetry (DSC)[136] are some of the recent techniques often used to characterize various viral capsids.

1.7.5. ProCharTS as a novel approach to study HBV Cp capsid assembly

The methods discussed above have some limitations as some are unable to detect heterogenous morphology, whereas others are high in cost, labour intensive, or invasive in nature, causing damage to samples. Thus, we try to explore the intrinsic optical properties of the HBV capsids and dimers through our novel technique, ProCharTS and analyse the aggregation kinetics by ProCharTS absorbance/luminescence. It is a simple method and requires only a portable UV Vis spectrophotometer or fluorimeter compared to the above techniques. It thus eliminates the need for careful pre-processing or advanced infrastructure and the high cost to characterise the viral assembly process. Emerging reports of Clustering Triggered Emission (CTE) from various common polymers and also from small molecules without aromatic amino acids displaying deep blue luminescence have been motivating. Viral capsid assembly, being considered a systematically controlled aggregation forming clusters of individual HBV core protein in an icosahedral symmetry, should enhance molecular 3D contacts within charged species. Such large clusters of HBV dimers in capsids are expected to give enhanced ProCharTS as opposed to when in isolated monomeric forms. It thus can reveal the details of molecular level contacts among key charged atoms and residues within the capsids. Moreover, it being a non-invasive technique is not expected to interfere with the capsid assembly process, unlike many fluorescent-based techniques.

Chapter 2

Experimental Techniques used; Materials and Methods

2.1. Experimental techniques

In this section, I discuss mainly the techniques used to accomplish the various experiments. UV-Visible Absorption Spectroscopy, Fluorescence spectroscopy, Circular Dichroism, Light Scattering, and Size Exclusion Chromatography are the important techniques that broadly encompass the thesis work and are explained briefly under various headings.

2.1.1. Absorption spectroscopy

Absorbance is the measure of the quantity of light absorbed by a sample using spectroscopy, particularly for quantitative analysis. The general setup of a spectrometer is shown below in **Figure 2.1**

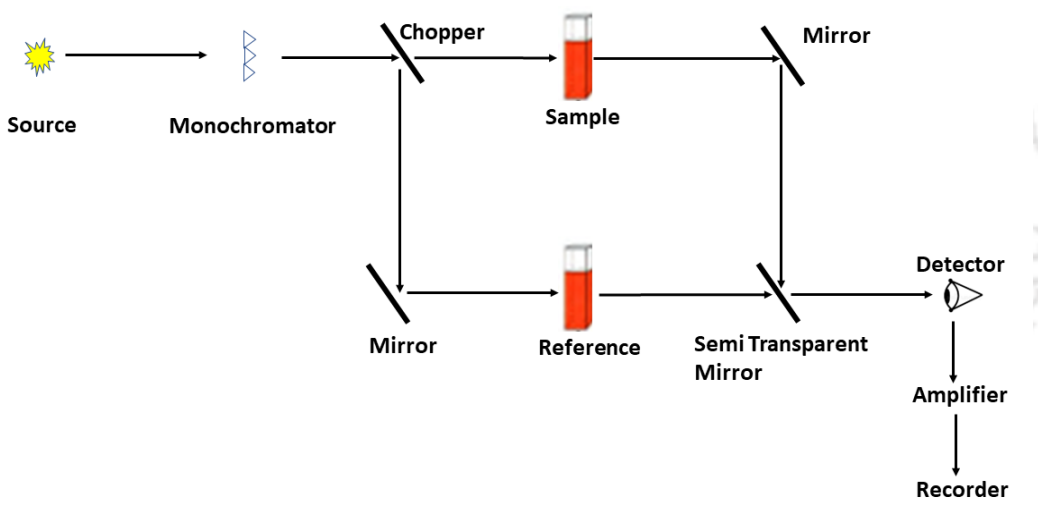


Figure 2.1: Schematics of UV-Visible spectroscopic set up. (Adapted from Cantor & Schimmel, 1980)[137]

When the light with intensity I_o, incident on the sample in a section sufficiently thin section (dl) having total path length 'l' (in cm) with a molar extinction coefficient of ε (M^{-1} cm^{-1}) and concentration 'C', then the fraction of incident light absorbed by chromophores in the sample is given by

$$-\frac{dI}{I} = C\varepsilon' dl \qquad (2.1)$$

Integrated the equation 2.1, we get-

$$A(\lambda) = log\frac{I_o}{I} = C\varepsilon(\lambda)l \qquad (2.2)$$

where ε = ε′/2.303 and *A* is the optical density or absorbance of the sample.

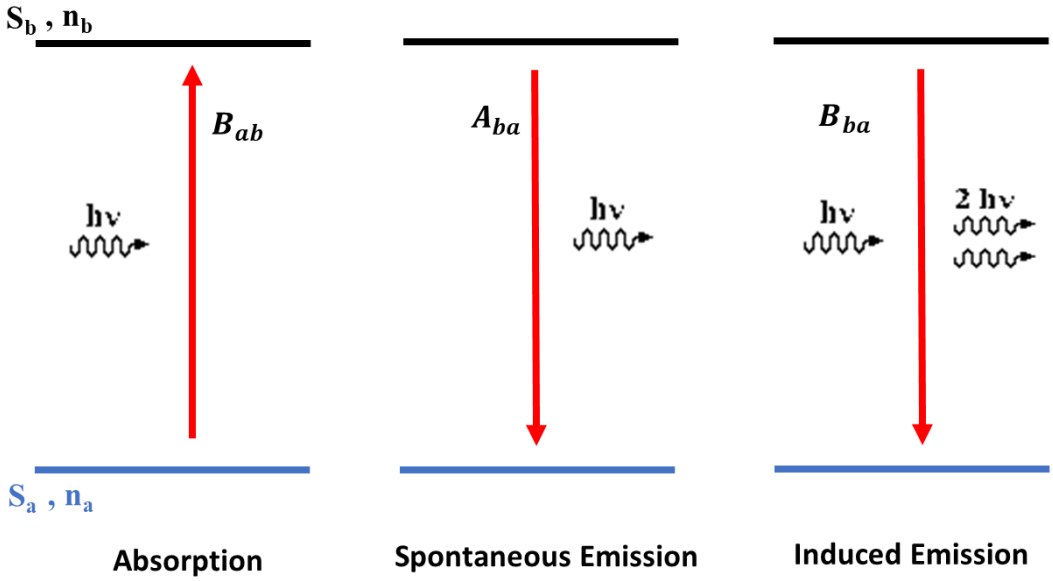


Figure 2.2: Schematics of the mechanism for absorption, spontaneous emission and induced emission.

Electronic transitions take place in femtosecond timescale from the singlet ground state (S_a) to the various higher excited states (S_b)[2,137]. Einstein coefficient for stimulated absorption is an essential parameter in absorption which measures the probability of absorption by an electron from state a with n_a density of electrons to state b with n_b density of electrons as shown in **Figure 2.2** and is represented by-

$$\frac{dn_b}{dt} = -\frac{dn_a}{dt} = B_{ab}n_a\rho(\nu) \qquad (2.3)$$

B_{ab} is Einstein's coefficient for induced absorption, which is related to the molar extinction coefficient by the following equation

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

where N_0 is Avogadro's number, ν is the frequency of light, c is the velocity of the light and $\rho(\nu)$ represents the energy density of the isotropic radiation field for the transition.

Two other important parameters are used to define the absorption phenomenon. D_{ab} is the dipole strength, which measures the transition dipole strength, and f_{ab} is the oscillator strength that expresses the probability of absorption that compares the absorption intensity to that of the expected quantum mechanical transition rate. D_{ab} and f_{ab} are related to the molar extinction coefficient by-

$$f_{ab} = (8\pi^2 mc/3h\nu)D_{ab} = 4.315x10^{-9}\int(\varepsilon'/\nu)\,d\nu \qquad (2.5)$$

where m is the mass of an electron.

2.1.2. Emission

Subsequent to absorption, the electrons from the higher excited state (S_b) come down to the singlet ground state (S_a), emitting photons. This process is called emission. Two types of emissions can broadly occur, as seen in **Figure 2.2.** First, through spontaneous emission, the electrons can come down to the ground state emitting photons. Secondly, the electrons, coming down to the ground state, encounter another photon and emit two photons of the same frequency. The rate of emission from the excited state (S_b) to the singlet ground state (S_a) is given by-

$$\frac{dn_a}{dt} = -\frac{dn_b}{dt} = A_{ba}n_b \;\; (spontaneous) \qquad (2.6)$$

$$\frac{dn_a}{dt} = -\frac{dn_b}{dt} = B_{ba}n_b\rho(\nu) \;\; (induced) \qquad (2.7)$$

where A_{ba} and B_{ba} are the Einstein's coefficient for spontaneous and induced emission, respectively.

The relationship between Einstein's coefficient for spontaneous A_{ba} and induced B_{ba} are the emission can be expressed as-

$$\frac{A_{ba}}{B_{ba}} = \frac{8\pi h\nu^3}{c^3} \qquad (2.8)$$

where ν is the frequency of light and c is the speed of light.

Thus, the higher frequency of light, the more is the probability of spontaneous emission. Spontaneous emission is responsible for most of the luminescence and Fluorescence we observe in the real world, which is discussed in the next section.

2.1.3. Fluorescence

Fluorescence spectroscopy is one of the most routinely used techniques in biochemistry and biophysics due to its sensitivity to detecting the dynamic fluctuations in proteins and other macromolecules. The sensitivity arises since the emission process occurs at a nanosecond, and any perturbations or effect of solvent interacting with the fluorophore is reflected in the emission spectrum. Fluorescence is mediated by electronic transitions from excited singlet state (S_1) to ground singlet state (S_0) with visible or ultraviolet light emission. Emitted light generally has a longer wavelength than absorption wavelength due to the loss in energy through various non-radiative processes, as shown in **Figure 2.3.** The non-radiative process includes the A) Internal Conversion due to vibrational relaxation from the higher S_2 states to the S_1 state in the picosecond timescale, dissipating heat. B) Intersystem crossing where an excited electron in the S_1 state can spin convert itself to the first excited triplet state T_1. The electron from the T_1 state then comes to the S_0 state at a longer time scale (in milliseconds) since the transition is spin forbidden. C) Quenching may occur by either depopulating the fluorophore in the excited state or interacting with the ground state fluorophore forming a non-fluorescent complex. All these three processes significantly affect the fluorescence intensity of a chromophore.

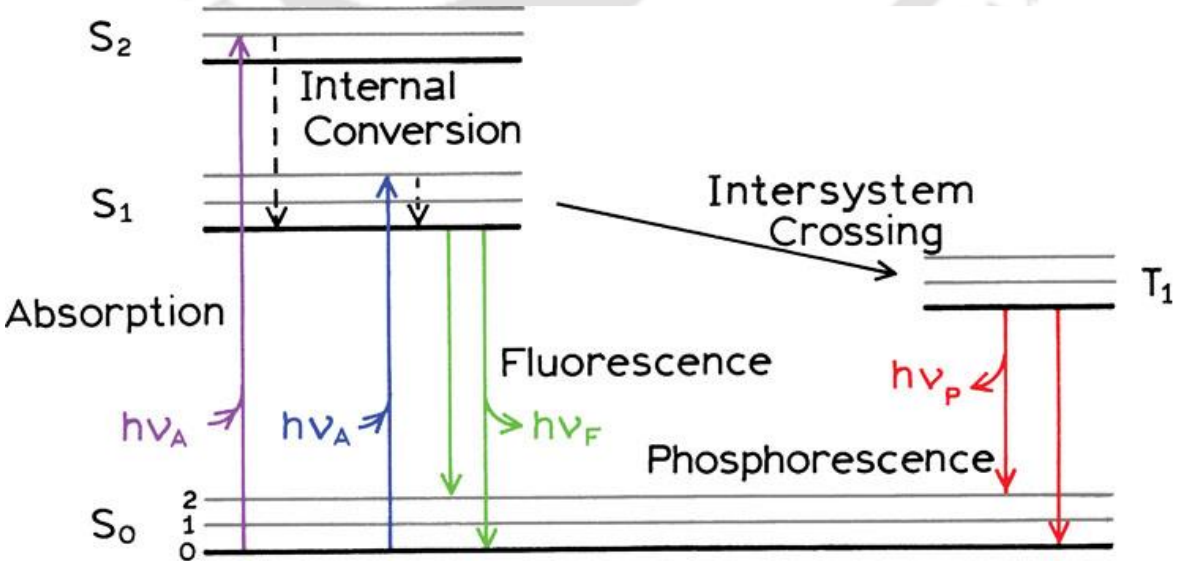


Figure 2.3. A simplified form of the Jablonski diagram. (Adapted from Lakowicz, 2006)[2]

For a fluorophore where both the non-radiative as well as the radiative process occurs, the rate of Fluorescence significantly decreases, and the observed fluorescence lifetime (τ) is expressed as-

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

where k_r, k_{isc} and k_{ic} represent the rate of the radiative decay, intersystem crossing, internal conversion. k_q and $[Q]$ represent the fluorescence quenching rate and molar quencher concentration. The above equation can be simplified by combining all the rates of non-radiative decay processes as k_{nr}.

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

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

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

The natural lifetime (τ_n) of a fluorophore is defined as the case where no non-radiative process occurs and can be represented as-

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

Combining equations 2.10, 2.11 and 2.12, we get-

$$\Phi = \frac{\tau_F}{\tau_n} \qquad (2.13)$$

This equation represents that quantum yield decreases with decreases in fluorescence intensity caused by non-radiative processes.

2.1.3.1. Steady-state fluorescence

Fluorescence measurements are broadly classified into two types: steady-state and time-resolved measurements. Here, the steady-state Fluorescence is discussed. In general, when light is incident on a fluorophore, the excited electrons jump from the S_0 state to the S_1 state. The

electrons then come down from the lowest vibrational level of S_1 to the ground state, emitting Fluorescence. In steady-state, an equilibrium is created between the population of electrons in the excited states S_1 and ground-state level S_0. The fluorescence intensity arising from the equilibrium state is measured as a function of wavelength and used to calculate various other fluorescent properties: quantum yield of a fluorophore, steady-state anisotropy, Fòrster resonance energy transfer, static or dynamic quenching etc. In contrast to time-resolved Fluorescence, steady-state Fluorescence is limited to time-averaged information of excited state fluorophore. Steady-state Fluorescence thus can reveal various information about the local fluctuations and perturbations around any fluorophore.

2.1.3.2. Steady-state fluorescence anisotropy

Steady-state Fluorescence anisotropy is an essential aspect of fluorescence measurements which can give important information regarding the shape and size of the fluorophore by calculating the magnitude of depolarization through rotational motion. As shown in **Figure 2.4**, when vertically polarized light (along Z-axis) is incident on a fluorophore, the molecules are excited, and the extent of the absorption mainly depends on $cos^2\theta$, where the angle (θ) between the absorption dipole of the molecule and the electric vector of the light incident on it. This phenomenon is called photo selection. Absorption probability is maximum when both

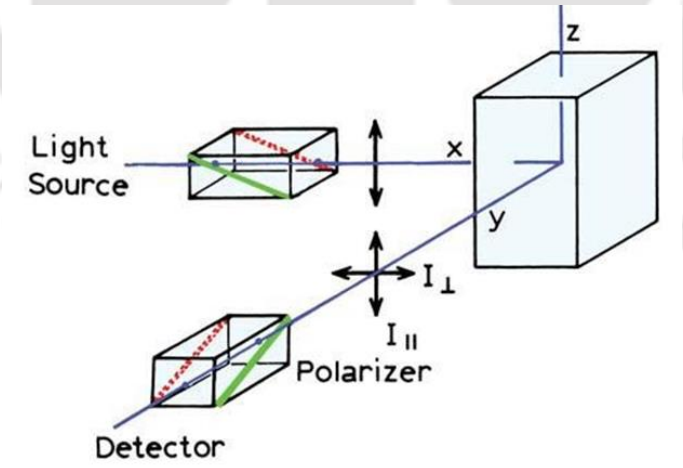


Figure 2.4: Schematic of the principle behind fluorescence anisotropy (Adapted from Lakowicz, 2006)[2]

the absorption dipole and electron vector are parallel (θ = 0°) and $cos^2\theta$ = 1 is maximum. Subsequently, the excited state fluorophores depolarize due to rotational diffusion in a similar timescale of fluorescence emission (in nanoseconds). The degree to which the depolarization is affected due to tumbling motion in the molecules is measured by the detectors with emission

polarizer either in parallel $I_{\parallel}$ or in perpendicular $I_{\perp}$ to the excitation polarizer in terms of steady-state anisotropy (r_{ss})

The following equation defines anisotropy-

$$r_{ss} = \frac{I_{\parallel} - GI_{\perp}}{I_{\parallel} + 2GI_{\perp}} \tag{2.14}$$

$$G = \frac{I_{HH}}{I_{HV}} \tag{2.15}$$

Here, G is a geometric factor defined by the ratio of the monochromator and detector sensitivity to the vertical I_{HV} and horizontal I_{HH} components of the polarized light.

Steady-state anisotropy (r_{ss}) can be explained in terms of the average of anisotropy decay over fluorophore fluorescence intensity decay as-

$$r_{ss} = \frac{\int_0^{\infty} r(t)\, I(t) dt}{\int_0^{\infty} I(t) dt} \tag{2.16}$$

Both intrinsic or extrinsic fluorophores attached to macromolecules can show steady-state anisotropy as long as the fluorophore has a fluorescence lifetime comparable to its rotational tumbling time. The anisotropy value is close to zero for smaller molecules with faster rotational diffusion. Moreover, anisotropy is not much sensitive where the intensity decay time of the fluorophore is much shorter than the rotational tumbling of the large macromolecule. Anisotropy has been successfully utilized to study various protein-protein; protein-ligand binding, protein aggregation, protein-DNA or protein-lipid membrane interactions.

2.1.3.3. Time-resolved fluorescence decay

Time-resolved measurements are generally performed either in the time domain or frequency domain. Here only the time domain fluorescence intensity decay is discussed. An ultrashort pulse (~10^{-10}-10^{-12} s) of excitation light far shorter than the decay time of the fluorophore of interest is incident on the sample. As soon as the excitation happens, the excited electrons from the S_1 state come down to the S_0 state emitting photon, and its fluorescence decay time is measured through a polarizer at 54.7° from the vertical Z-axis. This technique is called Time-

Correlated Single Photon Counting (TCSPC), and the set is shown in **Figure 2.5**. Less than one photon for each excitation is detected by keeping the fluorescence count rate at < 1% of the pulse repetition rate. The decay time is measured by the difference in arrival time of excitation pulse and emission through two different constant fraction discriminators (CFD). This decay time is measured in terms of increase in voltage ramp, through time to amplitude converter (TAC) and subsequently converted to a numerical value by a programmable gain amplifier (PGA). The process is done for N repetitions of photons arriving at the detector, and a photon distribution is obtained, representing the decay kinetics of the fluorophore. The decay kinetics is expressed as the slope of *log I(t)* vs *t* (ns), and the fluorescence decay time (τ) is measured.

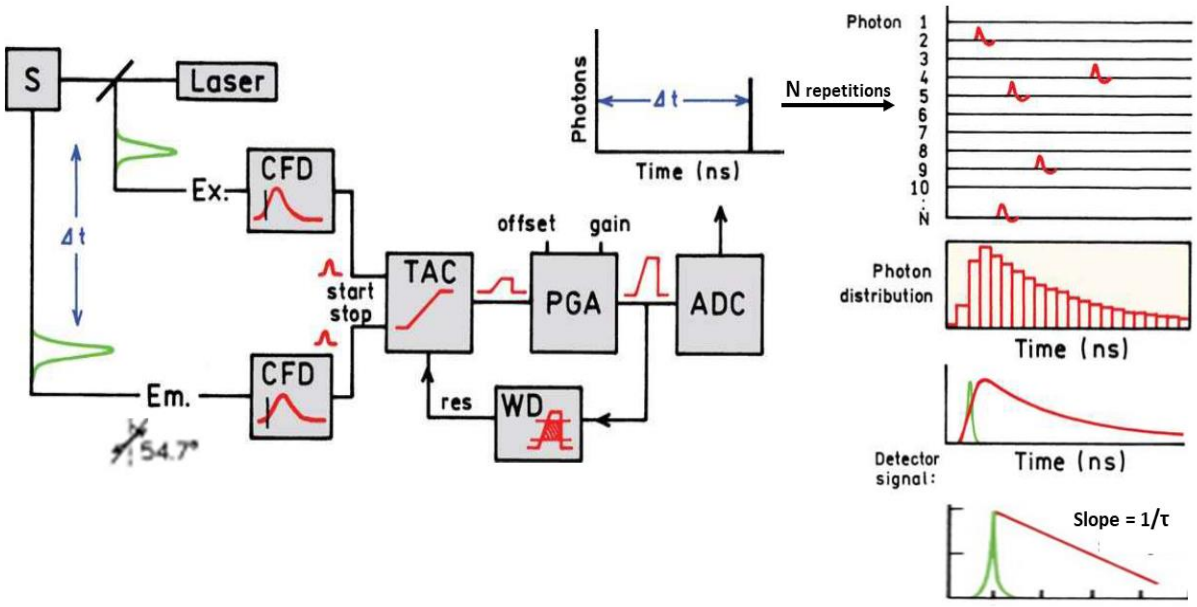


Figure 2.5: Electronic schematic for TCSPC. Adapted from (Lakowicz, 2006)[2]

The expression for the fluorescence intensity decay is expressed as-

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

where I_0 and $I(t)$ are the fluorescence intensity at time = 0 and t, respectively, and τ is the fluorescence lifetime of the fluorophore.

The average time a fluorophore stays in the excited state before returning to the ground state emitting a photon is represented by $<t>$. For a single exponential decay, this average time $<t>$ can be expressed as the average time t over the fluorescence intensity decay by-

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

Solving the equation,

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

Thus, for a single exponential decay, the average time of a fluorophore stay in the excited state equals fluorescence lifetime τ.

Fluorophores generally are not restricted to single exponential decay, and a single fluorophore may exhibit two, three or even four exponential decay components. For multi-exponential decay, the time-dependent intensity decay can be expressed as

$$I(t) = \sum_{i=1}^{n} \alpha_i \exp(-t/\tau_i) \qquad (2.20)$$

where τ_i and α_i represent the lifetime and amplitude of the i^{th} decay component, respectively.

$$\sum_{i=1}^{n} \alpha_i = 1 \qquad (2.21)$$

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

where τ_m is the mean lifetime of the fluorophore.

2.1.3.4. Time-resolved intensity decay analysis

The intensity decay curve essentially has three different curves: Instrument Response Function (IRF) $L(t_k)$, measured intensity decay $N(t_k)$ and calculated (fitted) intensity decay $Nc(t_k)$. The instrument response function (IRF) $L(t_k)$ is an instrument response to a sample with zero fluorescence lifetime, meaning the shortest time profile that TCSPC can detect. $N(t_k)$ is obtained by convoluting the actual intensity decay of the sample, $Nc(t_k)$, with the IRF $L(t_k)$. Each excitation impulse response $I_k(t)$ is a delta function excited at time t_k, which can be represented by-

$$I_k(t) = L(t_k) I(t - t_k)\Delta t \qquad (t > t_k) \qquad (2.23)$$

Since no emission is observed before time t_k, the term $(t- t_k)$ is used in the equation where the counted photons are collected into channels, each with width (Δt).

The measured intensity decay $N(t_k)$ can thus be expressed as the summation of each of the impulse responses as shown in **Figure 2.6** and is represented by-

$$N(t_k) = \sum_{t=0}^{t=t_k} L\,(t_k)\mathrm{I}(\mathrm{t} - t_k)\,\Delta\mathrm{t} \qquad (2.24)$$

Generally, the value of Δt is often small, and thus the above equation can be re-written as-

$$N(t) = \int_0^t \mathrm{L}(\mathrm{t}')\mathrm{I}(\mathrm{t} - \mathrm{t}')\mathrm{dt}' \qquad (2.25)$$

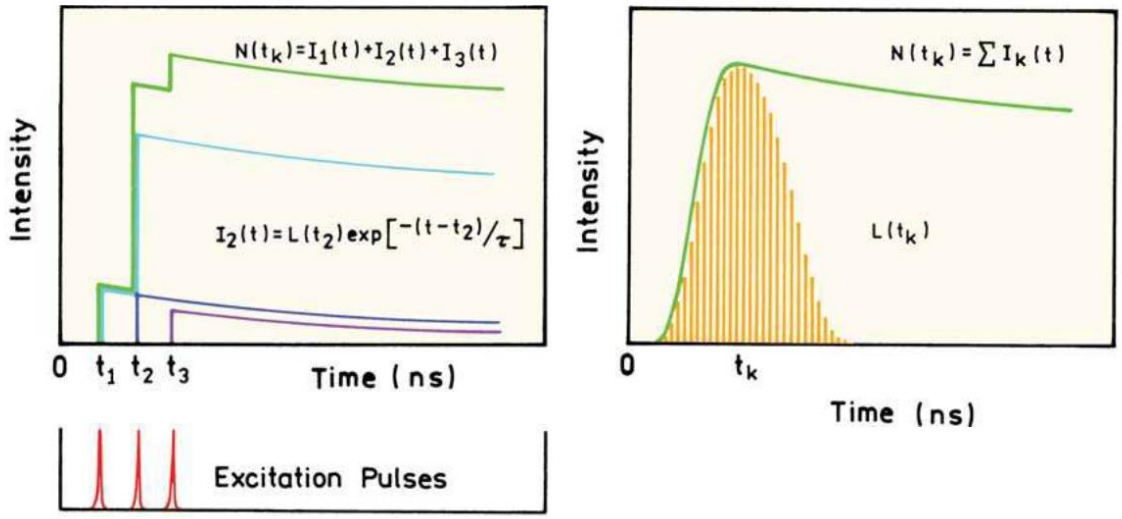


Figure 2.6: Convolution of an impulse response function I(t) with three excitation pulses $L(t_k)$ to yield the measured data $N(t_k)$. (Adapted from Lakowicz, 2006)[2]

The decay analysis for Time-resolved fluorescence data collected through TCSPC are generally analyzed by various methods: non-linear least square analysis (NLLS); maximum entropy method (MEM); Laplace transformation and phase-plane method etc. In this section, the analysis of the TCSPC data through Least square fit and maximum entropy analysis are discussed briefly.

In this NLLS method, the best match of the fitted decay $Nc(t_k)$, from measured decay $N(t_k)$, is obtained by minimizing the value parameter χ^2 to judge the goodness of fit. The χ^2 can be expressed as

$$\chi^2 = \sum_{k=1}^{n} \frac{1}{\sigma_k^2} [N(t_k) - Nc(t_k)]^2 = \sum_{k=1}^{n} \frac{[N(t_k) - Nc(t_k)]^2}{N(t_k)} \quad (2.26)$$

where σ_k is the standard deviation at each data point and n is the total number of channels used for the analysis.

Another more useful parameter χ_R^2 is often used instead of χ^2 which often eliminates the dependence on the number of data points' n' and is represented as-

$$\chi_R^2 = \frac{\chi^2}{n-p} = \frac{\chi^2}{v} \quad (2.27)$$

where p denotes the number of floating parameters and υ is the N degrees of freedom.

A value of χ_R^2 will be close to 1, signifies a better fit, and the iterative re-convolution process is carried out until χ_R^2 reaches close to 1.

Although χ_R^2 hints at the goodness of a decay fit, but it is not always necessary that χ_R^2 reaches close to 1 justifies the fitting in every possible scenario. For this, another parameter known as the distribution of residuals (D_k), which is essentially the difference between measured and fitted function, often supplements the χ_R^2 . The randomness of residual distribution signifies a better fit as compared to improper decay fit, as seen by some patterns in the distribution. The deviation (D_k) can be expressed as

$$D_k = \frac{I(t_k) - I_C(t_k)}{\sqrt{I(t_k)}} \quad (2.28)$$

Although the luminescence decay parameters are obtained by discrete non-linear least square analysis, there are two major inherent limitations related to the method. First, it assumes a model prior to the data set being fitted. Secondly, even if the decays are fitted assuming a model, for higher-order exponential decays, each model has to be fitted gradually from lower to higher-order exponential models to find out the model in which the decay will fit, which is quite tedious. In maximum entropy method (MEM) analysis[138,139], these limitations are overcome as it is based on a model-free approach, and no knowledge regarding the decay is required prior to fit. Each iteration is performed to generate a distribution of lifetimes by

maximizing the entropy S (known as Shannon-Jaynes entropy) and minimizing χ^2. The goodness of fit is characterized by χ^2 value as well as the increased S value, both of which are expressed as-

$$S = \sum_i -p_i \log p_i \qquad (2.29)$$

$$\chi^2 = \frac{(1/m)\ \sum_i^m [F_c\ (t_i) - F_e\ (t_i)]^2}{F_e\ (t_i)} \qquad (2.30)$$

where p_i is the probability for the i^{th} lifetime, m is the number of data sets and $F_c\ (t_i)$ and $F_e\ (t_i)$ are the calculated and experimental intensity decay, respectively, at the time t_i.

2.1.4. Circular Dichroism (CD)

Electromagnetic radiation comprises two oscillating components: electric and magnetic fields, which are perpendicular to each other and also to the direction of propagation. An oscillating electric field vector characterizes linear polarized light in only one plane where the direction of the vector stays constant, but its magnitude oscillates. Whereas, in circularly polarized light, the direction of the electric field vector continuously oscillates about its propagation axis, maintaining a constant magnitude, as shown in **Figure 2.7**. The prerequisite for displaying the CD spectrum is that the chromophore must have chirality since circularly polarized light by itself is "chiral". When circularly polarized light passes through a chiral centre, the speed, wavelength and absorptivity between right and left polarizations differ. Thus, the change in the molar circular dichroism depends not only on wavelength but also on the conformation of the absorbing molecule.

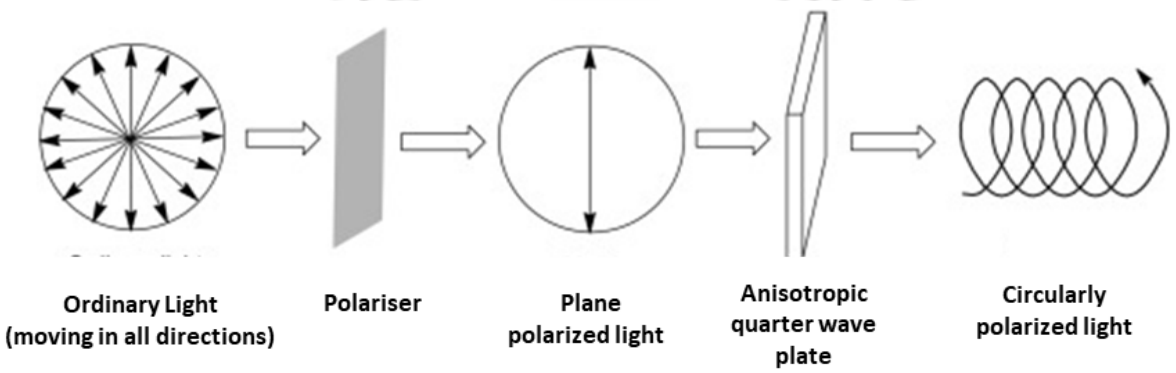


Figure 2.7: Origin of Circularly Polarized Light from ordinary light. (Adapted from Karnik & Hasan, 2021)[140]

Moreover, a linear charge displacement is observed in the electric field of light, whereas the magnetic field experiences a charge circulation. This causes the excitation of an electron in a helical motion, which both translational and rotational operators characterize. These quantities together reflect the rotational strength (R_{0a}) for state 0 to state a transition of the molecule. It is equivalent to the dipole strength measured in absorption. The relationship between the rotational strength and the molar ellipticity of a molecule [θ], an intrinsic parameter for polarization ellipticity, is given by

$$R_{0a} = \frac{3hc}{8\pi^3 N_0} \int \{[\theta(\lambda)]/\lambda\} d\lambda \qquad (2.31)$$

An optically active chiral molecule absorbs the right and left-handed circularly-polarized components of circularly polarized light in different amounts. Differences in absorbance can be represented as

$$\Delta A = A_L - A_R \qquad (2.32)$$

where A_L and A_R are the absorbance for left and right-handed circularly polarized light, respectively.

Molar absorptivity ΔA can be expressed in terms of molar circular dichroism $\Delta\varepsilon$ using Beer-Lambert Law $(A = \varepsilon Cl)$

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \qquad (2.33)$$

$$\Delta A = \Delta\varepsilon Cl \qquad (2.34)$$

The CD spectrum is mostly reported in terms of ellipticity, θ, which is, in fact, a measure of the polarization ellipticity. It can be represented in terms of the electric field as given by-

$$tan\,\theta = \frac{E_L - E_R}{E_L + E_R} \qquad (2.35)$$

where E is the magnitude of the electric field vector of right and left circularly polarized light

The ellipticity can be written in terms of molar absorptivity as-

$$\theta(degrees) = \Delta A \left(\frac{ln10}{4}\right)\left(\frac{180}{\pi}\right) \qquad (2.36)$$

Moreover, similar to molar circular dichroism, molar ellipticity defines the intrinsic property of the chiral molecule by removing the dependence of concentration and pathlength and is given as,

$$[\theta] = \frac{100\,\theta}{Cl} \qquad (2.37)$$

Combining equations 2.34, 2.36 and 2.37, we get the relationship between molar ellipticity and molar dichroism,

$$[\theta] = 3300\Delta\varepsilon \qquad (2.38)$$

The units of molar ellipticity are historically (deg. cm^2/dmol).

Circular Dichroism (CD) is used to detect the changes in conformations in proteins and nucleic acids. The far-UV CD spectra in the region of 240 to 190 nm give information about the secondary structure of the protein. In contrast, Near-UV CD spectra in the region of 250 to 290 nm reveal information about the tertiary structure of the protein. When the chromophores in the amides of the polypeptide chain in proteins are aligned in a specific orientation, the optical transitions experiences "exciton" interactions and are split into multiple transitions. This is why different secondary structures like α-helix, β-sheet, β-turn and random coil show characteristic individual CD signals.

2.1.5. Light scattering

Light Scattering is an important phenomenon used to gain information about biomolecules in solution. As compared with X-Ray Diffraction studies where the static orientation of biopolymers is studied from crystallized samples, light Scattering reveals the actual conformation of the molecules in solutions since most of the biomolecules are soluble rather than crystallized. When light is incident on macromolecules in solution, the light is scattered in all directions with varying intensity and is detected by the detector. Information regarding

the size and shape of macromolecules is stored in the scattered light[141]. Here we discuss two types of Scattering: static light Scattering and dynamic light scattering.

Static light scattering (SLS) is a technique to measure absolute molecular weight by measuring scattered light intensity as time-averaged intensity giving information regarding molecular weight and radius of gyration of macromolecules. For a system containing a biomolecule with molecular weight *M* having a concentration of *C*, the equation for Rayleigh scattering can be written in terms of the Rayleigh ratio R_θ as-

$$R_\theta = \frac{2\pi^2 n_o^2 (dn/dC)^2}{N\lambda^4} CM \qquad (2.39)$$

$$R_\theta = \frac{i_\theta r^2}{I_o(1 + cos^2\theta)} \qquad (2.40)$$

where n_o is the refractive index of the solution, N is the Avogadro's number, and λ is the wavelength of the incident light

The measurement of R_θ is generally carried out in a photometer, and the intensity of the light scattered is measured at a 90° angle with respect to the incident light and the molecular weight, M is estimated. Moreover, the molecule's radius of gyration (R_G) can also be estimated from the angular-dependent scattering experiments, considering the size of the molecule is large enough. The equation can be written as-

$$P_\theta = 1 - \frac{h^2 R_G^2}{3} + \cdots \qquad (2.41)$$

where P_θ is the function defining how Scattering at an angle θ would be for a biomolecule if the molecule was shrunk to much smaller than λ. It should also be noted that this angular dependence of scattering relies explicitly on its shape only if the biomolecules are large enough. For small molecules, low values of θ or for long incident wavelength, P_θ approaches 1.

Dynamic Light Scattering (DLS) is used to characterize the size of protein aggregates, nucleic acids and other biopolymers based on their diffusion behaviour in solution. The fluctuation in the intensity of scattered light (due to Brownian motion) is defined by the translational

diffusion coefficient (D_τ) of the macromolecule in solution. The distribution in the size of the macromolecules is generated from the time autocorrelation function of the scattered light. The relation between the first-order autocorrelation function of scattered light and D_τ can be expressed as

$$G_\tau = 1 + \beta e^{-2D_\tau q^2 \tau} \qquad (2.42)$$

$$|q| = \frac{4\pi\eta_o}{\lambda_o} \sin(\theta/2) \qquad (2.43)$$

Here, β is a parameter which depends upon the geometry and optical setup of the instrument, τ is the decay time, q is a scattering vector, η_0 is the refractive index, λ_0 is the wavelength of incident light in vacuum, and θ is the scattering angle.

The relation between D_τ and the hydrodynamic radius (R_h) of the spherical macromolecule can be expressed by the Stokes-Einstein equation as:

$$D_\tau = \frac{k_B T}{6\pi\eta R_h} \qquad (2.44)$$

where k_B is Boltzmann's constant (1.381 x 10^{-23} J/K), T is absolute temperature, η is the viscosity of the medium and R_h is the hydrodynamic radius of spherical macromolecule.

2.1.6. Size Exclusion Chromatography

Size exclusion chromatography (SEC) is mainly used to separate higher molecular-weight species, such as polymer, large proteins or aggregates, from lower molecular weight proteins. Its principle is based on the movement of the molecules based on their size differences. When a sample containing various-sized solutes is injected into the column, the individual solutes, based on their size, travel different paths and are carried by the solvent through the highly rigid and porous material of the packed column, as shown in **Figure 2.8**. The smallest molecules can penetrate deeply into smaller pores of the column, whereas the larger molecules cannot enter such smaller pores. Small solutes travel longer paths and thus elute at longer times (more retention time) than larger solutes. In both liquid and gas chromatography, the movement of

the solute is often governed by the type of mobile phase as it is held back either by the column packing or the stationary phase. However, in Size Exclusion Chromatography, molecular size or the hydrodynamic radius dictates the separation of the solutes.

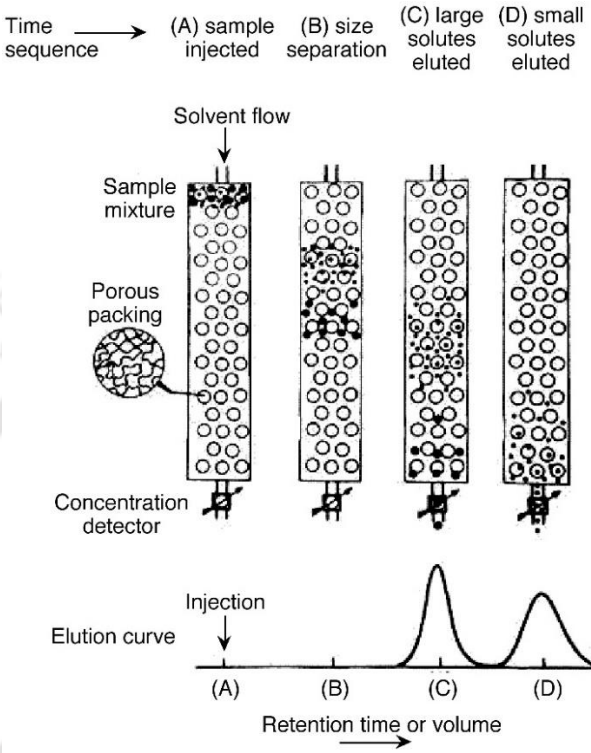


Figure 2.8: The schematic of the principle of Size Exclusion Chromatography (Adapted from Barron, 2015)[142]

The molecular weight (M_w) of a solute is directly linked with the retention time or the elution volume. M_w can be estimated by plotting a standard curve between partition coefficient (K_{av}) against the log (M_w) of standard proteins. K_{av} can be expressed in terms of elution volume

$$K_{av} = \frac{V_e - V_o}{V_t - V_o} \qquad (2.45)$$

V_e, V_o, and V_t are the measured elution volume, void volume and the total column volume.

For polymeric biopolymers, the weight averaged molecular weight (M_w), and the number averaged molecular weight (M_n) composed of N_i numbers of individual solutes with sizes of M_i can be defined by the

$$M_n = \frac{\sum N_i M_i}{\sum N_i} \qquad (2.46)$$

$$M_w = \frac{\sum N_i M_i^2}{\sum N_i M_i} \qquad (2.47)$$

The polydispersity index (PDI), a significant parameter in polymer chemistry, can be estimated by the M_w / M_n ratio. It measures the distribution of the molecular mass of individual chains in a biopolymer. The more M_w / M_n is, the more highly dispersed the polymer is. A ratio nearer to 1 corresponds to a perfectly monodisperse polymer.

2.2. Materials

2.2.1. Chemicals and Reagents

Glycine (G8898); Streptomycin Sulfate salt (S6501); Triton X-100 (T9284); β Mercaptoethanol (63689); PMSF (P7626); Sinapic acid (85429); $CaCl_2$ (C8106); Urea(U5378); 1, 4-Dithiothreitol (D0632); Folin & Ciocalteu's phenol reagent (F9252); Guanidine hydrochloride (G3272); Gel loading buffer (G2526); Trifluoroacetic acid (T6508); Thioflavin T (T3516); 8-Anilinonaphthalene-1-sulfonic acid (A3125); Acrylamide (A3553); L-Arginine (A5006); L-Arginine monohydrochloride(A5131); L-Aspartic acid potassium salt (A6558); L-Histidine (53319); poly-L-arginine hydrochloride(P4663); poly-(α,β)-D,L-aspartic acid sodium salt (P3418); N-acetyl tryptophan amide (A6501) were purchased from Sigma Aldrich Chemicals Pvt. Limited, Bengaluru, India. The Luria Broth (M1245); Luria Agar (M557); $MgCl_2$ (TC186) and Imidazole (GRM1864); Lithium Chloride (GRM768); HEPES (MB016); Dithiothreitol (RM525); Kanamycin sulphate (MB105); Isopropyl β-D-1-thiogalactopyranoside (MB072); Sodium chloride (GRM031); Sodium dihydrogen Phosphate monohydrate (PCT0013); Ampicillin sodium salt (TC021); EDTA(GRM3915); Sucrose (MB025) were procured from HiMedia Laboratories, India. 5-(Dimethylamino) naphthalene-1-sulfonyl chloride (D23) was purchased from Invitrogen, Molecular Probes, USA. The PD-10 desalting column (17-0851-01) was purchased from GE Healthcare. The Tris (Hydroxy methyl amino methane) (93315); SDS (184190); and Potassium chloride (91635); Cyclohexane (1028220500); Ammonium Sulfate (101217); Sodium hydrogen carbonate (106329); Sodium hydroxide (193102); Dimethylformamide (17754) were purchased from Merck India Limited. Nuvia IMAC Ni-Charged Resin (780-0800) was purchased from Bio-Rad. All chemicals used were analytical grade with purity ≥98%.

2.2.2. Amino acid sequences of the proteins

A. Symfoil 4P/PV2 (131 amino acids) [PDB ID: 4D8H]

HHHHHHEVLLRSTETGQLLRINPDGTVDGTRDRSDPGIQLQISPEGNGEVLLRSTE
TGQLLRINPDGTVDGTRDRSDPGIQLQISPEGNGEVLLRSTETGQLLRINPDGTVDG
TRDRSDPGIQLQISPEGNG

B. Hepatitis B Virus Core protein (149 amino acids) [PDB ID: 1QGT]

MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQ
AILCWGELMTLATWVGNNLEDPASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRET
VLEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVV

Charged amino acids: Arginine(red); Aspartate(blue); Lysine(green); Glutamate(purple); Histidine(green) are highlighted in bold fonts.

2.2.3. Amino acid sequences of the switch peptides

Switch peptides with various combinations of Lysine, Glutamate residues with the core fragment Aβ 16-22 (KLVFFAE) intact are shown below-

Pep1[2Ac] ***(1K/1E)*****:Ac-SLKLVFF(H$^+$)SELSG-NH$_2$**

Pep2 ***(1K/1E)*****: NH$_2$-SLKLVFF(H$^+$)SELSG-NH$_2$**

Pep3 ***(1K/0E)*****: NH$_2$-SLKLVFF(H$^+$)SSLSG-NH$_2$**

Pep4 ***(2K/0E)*****: NH$_2$-SLKLVFF(H$^+$)SKLSG-NH$_2$**

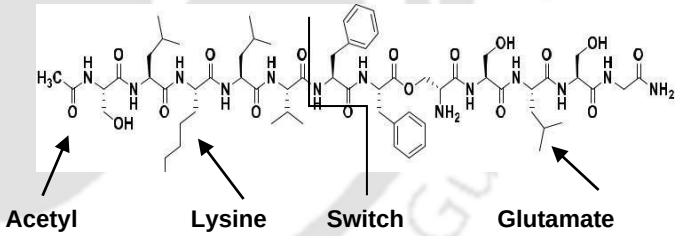
Pep5 ***(0K/1E)*****: NH$_2$-SLSLVFF(H$^+$)SELSG-NH$_2$**

Pep6 ***(0K/2E)*****: NH$_2$-SLELVFF(H$^+$)SELSG-NH$_2$**

Pep7 ***(0K/0E)*****: NH$_2$-SLSLVFF(H$^+$)SSLSG-NH$_2$**

Pep8[7Ac] ***(0K/0E)*****: Ac-SLSLVFF(H$^+$)SSLSG-NH$_2$**

Pep9[3Ac] ***(1K/0E)*****: Ac-SLKLVFF(H$^+$)SSLSG-NH$_2$**



Charged amino acids: Lysine (red); Glutamate (blue); acetylated N terminus (green) are highlighted

We used a peptide concentration of 50 μM for all peptides, but we could only observe controlled aggregation in peptides 1 and 9. In contrast, peptides 5, 7, 8 turned turbid and thus precipitated out of the solution, meaning aggregation assay could not be performed, therefore

discarded for this study. Consequently, we tried to increase the working concentration of the rest of the peptides to 200 μM. It turned out that peptides 2 and 3 both showed aggregation, as observed in Circular Dichroism data discussed later and is used for this case study. In contrast, pep 4 and pep 6 did not show any aggregation even in that condition and are not being studied here. Thus, we were left with four peptides, two of which were acetylated (peptides 1 and 9) with a working concentration of 50 μM, whereas both the non-acetylated (peptides 2 and 3) were kept at 200 μM. In addition to that, the aggregation mechanism of pep 1 with a working concentration of 200 μM was also studied here. Thus, to sum up, we have 5 cases named as mentioned below and used the same terminology everywhere in this article.

Pep I: Ac-SLKLVFF(H^+) SELSG-$CONH_2$ (50 μM)
Pep II: Ac-SLKLVFF(H^+) SSLSG-$CONH_2$ (50 μM)
Pep III: NH_3^+-SLKLVFF(H^+) SELSG-$CONH_2$ (200 μM)
Pep IV: NH_3^+-SLKLVFF(H^+) SSLSG-$CONH_2$ (200 μM)
Pep V: Ac-SLKLVFF(H^+) SELSG-$CONH_2$ (200 μM)

2.3. Methods

2.3.1. Protein(s) purification

2.3.1.1. Cloning

Symfoil 4P PV2 containing 131 amino acids were cloned in a pET21a(+) expression vector with a hexahistidine (His6) tag fused with terminus. The construct was provided by Prof. Michael Blaber from Florida State University, USA. On the other hand, HBV Cp 149 were cloned in pRSF_T7-HBc149optTetPrRep plasmid vector with a T7 promoter and had no additional tag attached to them. The construct was provided by Prof. Michael Nassal from the University of Freiburg, Germany.

2.3.1.2. Preparation of competent cells

E. coli BL-21 (DE3) and *E. coli* BL-21 codon$^+$ (DE3) strains were used for Symfoil 4P/PV2 and HBV Cp 149, respectively. For both strains, the method of preparing competent cells is the

same, which is described below. A single bacterial colony was collected from the plate of Luria Agar and inoculated into 5 mL of Luria Broth (LB) medium and allowed to incubate at 37 °C and 180 rpm for 10-12 hours in a shaking incubator. 500 µL of the culture was then collected and allowed to grow in a 50 mL volume of LB medium under similar conditions till the culture turned slight turbid and absorbance at 600nm(A_{600}) ~0.3 or 0.4 was reached. The culture was then placed on ice for 10-15 minutes, cells were centrifuged at 3000 rpm for 10 minutes at 4 °C, and the pellet was collected. The pelleted cells were then resuspended in a 15 mL solution of 80 mM $MgCl_2$ + 20 mM $CaCl_2$ cold solution and were centrifuged again at similar conditions. The pelleted cells thus collected were mixed slowly and gently in a solution mixture of 900 µL of 0.1 M $CaCl_2$ + 100 µL glycerol and were kept as aliquots of 100-200 µL in microcentrifuge tubes at -80 °C for further use.

2.3.1.3. Transformation

The *E. coli* BL-21 (DE3) and *E. coli* BL-21 codon$^+$ (DE3) competent cells kept at -80 °C were thawed, and about 0.2-0.6 µL of the Symfoil PV2 and HBC Cp plasmid respectively were added and incubated for 15 minutes at 4 °C. In each case, a negative control set with competent cells with no plasmid was also prepared. After incubation, the cells in microcentrifuge tubes were placed at 42 °C for 90 seconds and transferred to ice immediately. After a few minutes of incubation, the cells were suspended in 800-900 µL of LB media with no antibiotic and were incubated in a shaking incubator at 37 °C and 180 rpm for 1 hour. As soon as the culture turned slight turbid, they were centrifuged at 6,000 rpm for 10 minutes. The cells were then pelleted and resuspended in 100 µL of fresh LB media and plated on an LB agar plate with Ampicillin (100 µg/mL) for Symfoil 4P/PV2 and Kanamycin (50 µg/mL) for HBV Cp. The LB agar plates were then kept in an inverted manner in a static incubator at 37 °C for 10-12 hours until colonies of transformed cells started appearing.

2.3.1.4. Expression

E. coli BL21 cells transformed with individual recombinant plasmids were grown in LB broth containing 100 µg/mL ampicillin at 37°C. For transformed *E. coli* BL-21 codon$^+$ (DE3) cells, LB broth containing 50 µg/mL kanamycin at 37°C was used. When the culture achieved an optical density of 0.6–0.7 at 600 nm, Isopropylthio-β-d-galactoside (IPTG) was added to a final concentration of 1.0 mM for both the *E. coli* strains. After an additional 3-4 hours of incubation,

the cells were harvested by centrifugation at 6000 rpm for 15 min, and the pelleted cells were kept at -20°C.

2.3.1.5. Purification

Purification of Symfoil 4P/PV2 was carried out similar to that of previously reported protocol[143] with some modifications. From a 1litre culture, the pelleted cells were mixed in ice-cold lysis buffer (50 mM phosphate buffer (pH 7.4); 500 mM NaCl; 5 mM Imidazole). Subsequently, 2 µL of 20% Tween 20 and 1mM phenyl-methyl-sulfonyl-fluoride (PMSF) were added, and the cells were lysed by sonication. The cell lysate was then centrifuged for 30 mins at 13000g, 4°C. Supernatants were filtered with a 0.45 µm filter and were loaded on the affinity column of nickel–NTA agarose resin (Qiagen) with 2 mL Ni-NTA agarose beads. Before supernatants were loaded, the column was pre-equilibrated with binding buffer (50 mM phosphate buffer (pH 7.4); 500 mM NaCl; 5 mM Imidazole). The resin containing the protein was kept for 4-5 hours in a rocker at 4 °C allowing the NI-NTA beads to bind to His tag of protein. 2-3 column volumes (CVs) of the same binding buffer were used to wash the protein-bound-Ni-NTA beads. The proteins were then eluted with elution buffer I (50 mM phosphate buffer (pH 7.4); 500 mM NaCl; 50 mM Imidazole) and with elution buffer II (50 mM Phosphate buffer (pH 7.4); 500 mM NaCl; 250 mM Imidazole). The samples were collected in multiple fractions. Quantifying proteins in these fractions was done using UV-Visible spectroscopy, and fractions with relatively higher protein concentrations were pooled together. The protein was dialyzed against enough water to remove the imidazole from the elution buffer. Subsequently, the proteins were further characterized by SDS-PAGE and MALDI -TOF to check the purity of the purified proteins.

HBV Cp 149 purification was carried out similar to that reported previously[144] with minor modifications. Cells harvested from 1-litre culture were lysed and sonicated in 50 mM Tris-HCl buffer. 1 mM EDTA, 5 mM DTT, 1 mM PMSF, 0.025-0.05 mg/mL DNase (pH 7.4) and centrifuged for 1 hr at 13000g, 4°C. Solid sucrose was added to 0.15M to the supernatant and was ultracentrifuged at 100000g for 1 hr. Solid streptomycin sulphate was added to 1% w/v of supernatant and centrifuged for 30 min, 4°C. Subsequently, solid ammonium sulphate was added to 40% w/v of the supernatant and centrifuge for 1 hr, 4°C. The pellet is collected and stored at 4°C. The pellet was then dissolved in 100 mM Tris-HCl, pH 7.5,100 mM NaCl, 50 mM sucrose, and 2 mM DTT (purification buffer), filtered and passed through Hi-load

Superdex 16/600 200pg column (120 ml for one CV) at 1 mL/min during Size Exclusion Chromatography (SEC). The elute in the void volume was collected, which contained the HBV Cp 149 capsid. SEC was carried out using ÄKTApurifier^{TM} FPLC (GE Healthcare) chromatography system. Capsids collected were dialyzed in 50 mM Tris-HCl, pH 7.5, 0.15 M sucrose for 5-6 hours. The solution was titrated to pH 9.5 with 10 N NaOH, and solid urea was added to 3.5 M. The mixture was allowed to equilibrate for 1.5 hours at 4 °C to disrupt the capsid and form dimers. The dimers thus formed were then passed through the same SEC column equilibrated in 100 mM sodium bicarbonate, pH 9.5, 2 mM DTT. Pure HBV Cp 149 dimers were concentrated by Vivaspin centricon to 50 μM and stored in -80° C. All the SEC analyses were done after pre-equilibrating the column with 2 CVs, the same degassed buffer used for proteins.

2.3.2. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

To estimate the expression and purification of the proteins, reducing SDS-PAGE was performed based on the standard protocol[145]. Both the resolving gel (15% acrylamide) and stacking gel [a mixture of deionized water, Acrylamide, Tris(hydroxymethyl)aminomethane, SDS, Ammonium persulphate, N, N, N, N Tetra-methyl-ethylenediamine] were prepared accordingly for running all protein samples. Briefly, protein samples mixed in a 3:1 ratio with the gel loading buffer were heated to 95 °C for 3-5 minutes. Different samples (in separate wells); standard molecular markers to compare the approximate molecular weight of proteins (in one well) were added to the casted gels and allowed to be separated on vertical Mini-PROTEANR Tetra Electrophoresis System (Make: Bio-Rad). The system was run at 80-100 Volts for 1.5-2 hours until gel loading dye reached the bottom. Gels thus obtained were stained with Coomassie stain (prepared by dissolving 0.025% (w/v) Coomassie Blue R-250 in a solution of 45% methanol, 45% deionized water and 10% glacial acetic acid) for 4 hours. Subsequently, the gels were allowed to de-stain in a solution prepared by mixing 30% of methanol, 10% of glacial acetic acid and 60% of deionized water until clear bands of proteins and markers were visible.

2.3.3. Protein estimation

Both Symfoil PV2 and HBV Cp concentrations were estimated by the Lowry method[146]. For the standard curve, various dilutions of BSA ranging were prepared [0.05-1.00 mg/mL] from

BSA stock of 1.00 mg/mL dissolved in water. With BSA dilution, 200 µL of Symfoil PV2 and a blank (deionized water) were added to the aluminium foil-wrapped test tubes. Subsequently, 2 mL of ***Reagent I***: a solution prepared by mixing 48 mL of solution (2 % Sodium carbonate in 0.1 % NaOH, 100 mL; 1 mL of (solution 2.37 % Potassium sodium tartrate in water, 50 mL) and 1 mL of solution (1.56 % Copper sulfate in water, 50 mL) was added to each of the tubes and vortexed thoroughly. After all the samples were thoroughly mixed with the reagents, they were incubated at room temperature for 10-15 minutes. To each of the samples, 200 µL of ***Reagent II,*** prepared by mixing an equal volume of Folin's reagent and water, was added, mixed, and incubated in a dark space for 30 minutes. Finally, the absorbance of each sample was measured in duplicates at 650 nm wavelength. Using the standard BSA curve, the protein concentration was estimated quite accurately from the concentration versus absorbance plot. However, while performing the experiments for Aβ switch peptides aggregation, the lyophilized switch peptides were weighed, and their respective concentrations were calculated instead of the Lowry method.

Moreover, both concentration of Symfoil PV2 and all the Aβ switch peptides were further validated by measuring absorbances values 215 nm and 225 nm (A_{215} and A_{225}) respectively in the far UV region and using the following equation[147]

$$\text{Protein/peptide concentration (µg/ml)} = 144*(A_{215}-A_{225}) \qquad (2.48)$$

Concentration of HBV Cp149 was calculated by taking the absorbance in UV region and using extinction coefficient ε_{280} of 60800 $M^{-1}cm^{-1}$ for HBV Cp149 dimers[124].

2.3.4. Mass Spectrometry

MALDI-TOF mass spectrometer (Daltonics Bruker, Germany) was used to determine the mass of Symfoil PV2 and HBV Cp. A solution of Sinapic acid was dissolved in TA-30 solvent [0.1% Trifluoroacetic acid and acetonitrile mixed in 7:3 ratios] to prepare the MALDI matrix. It was then sonicated for 30 minutes, centrifuged at 12000 rpm at 25 °C for 15 minutes, and the supernatant collected carefully was mixed with protein in a 2:1 ratio in PCR tubes. The mixture was then spotted on the target plate and was allowed to dry for 3 hours. Data acquisition was made with flex control in Linear Mode, and analysis was carried out by flex Analysis software (make: Bruker Daltonics).

2.3.5. Switch peptide aggregation

Switch peptides aggregation were performed as mentioned[148,149]. All the switch peptides were weighed and dissolved in 0.1%(v/v) TFA (Trifluoroacetic acid) at pH 2, and the stock was made up to ~2 mM. From the stock, two final working concentrations of 50 μM and 200 μM were selected and the peptides were allowed to aggregate in phosphate buffer pH 7.2. Different time scales were observed for each peptide, and their characterization was done using various techniques at multiple time points.

2.3.6. HBV Cp capsid assembly

HBV Cp capsid assembly was performed as similar as reported previously[124]. The HBV Cp 149 dimer that was purified was dialyzed at 20 °C for 4 hrs against 50 mM HEPES 2 mM DTT, pH 7.5 and its concentration was found to be ~50 μM and was used immediately. Capsid assembly was studied by dissolving 10 μM of HBV Cp in 50 mM HEPES plus 0.25 M or 0.5M NaCl, respectively, to initiate reaction and reaction was allowed for 30 minutes. Moreover, the assembly for 5 μM HBV Cp in 50 mM HEPES plus 0.5 M NaCl was also performed.

2.3.7. UV-Visible and ProCharTS absorbance

Double beam Lambda-25 UV-Visible Spectrophotometer (Perkin Elmer, USA) was used to record absorbance. A quartz cuvette with a path length of 10 mm (Hellma; Z600210) was used for all measurements. Absorption spectra were acquired in duplicates between 250 nm and 800 nm with a scan speed of 480 nm/minute and are averaged. The spectra obtained from each concentration were normalized with individual concentration to calculate the extinction coefficient for all the samples

For amino acids, 10 different concentrations [0.1-1.0 M] for all amino acids except Histidine, where 6 concentrations [0.05-0.3 M] were used and their absorbance at different wavelengths were plotted versus concentration. The experiments were carried out in deionized water and at room temperature

For Arginine and Aspartate homo-polypeptides, 5 different concentrations: 0.5, 1.0, 2.5, 5 and 7.5 mg/mL were used for polypeptides and their absorbance at 7 different wavelengths was plotted versus concentration. The experiments were carried out in deionized water and at room temperature

For Symfoil PV2, absorbance was performed with 6 dilutions of protein: 15, 30, 50, 75, 100, and 125 μM and their absorbance at 8 different wavelengths was plotted versus concentration. Moreover, for quantifying Symfoil PV2, absorbance in far UV at 215 nm and 225 nm was also carried out by various sample dilutions. The experiments were carried out in deionized water and at room temperature

For switch peptide aggregation, ProCharTS spectra (250-800 nm) at different times for each peptide were carried out in aggregation buffer (phosphate buffer, 50 mM pH 7.2). The change in ProCharTS absorbance at various wavelengths versus aggregation time was calculated from each spectrum and plotted versus time. The peptide concentration was kept at 50 μM and 200 μM depending on the peptides. Moreover, to eliminate possible interference from Scattering, all the spectra were corrected for any possible scattering. The absorbance values were scatter corrected from 450-600 nm as log A vs log λ dependence and further extrapolating way till 250 nm[150]. All the readings were done in duplicates and at room temperature.

For HBV capsid assembly: Assembly of the 10 μM dimer in 50 mM HEPES with 0.25 M NaCl and 0.5 M NaCl, pH 7.4 (assembly condition) were monitored by collecting the UV-Visible spectrum from 250-800 nm for 30 minutes and was compared with dimer with the same concentration in 50 mM HEPES with no salt, pH 7.4 (non-assembly condition). From the spectra, absorbance at various wavelengths: 295, 310, 325, 340, 355 and 370 nm were plotted vs time (minutes) to monitor the kinetics.

Effect of amino acids on the NATA: NATA in the presence of a different concentration of amino acids were prepared. The concentration of NATA was kept constant at 20 μM in all the samples. A stock of NATA was prepared using an extinction coefficient of 5500 $M^{-1}cm^{-1}$ at 280 nm[151].

2.3.8. Steady-state luminescence

Luminescence Emission Spectra: Luminescence emission spectra for Symfoil PV2, homo-polypeptides and amino acids were measured in the range 300-540 (λ_{ex} 280); 315-570 (λ_{ex} 295); 330-600 (λ_{ex} 310); 360-660 (λ_{ex} 340); 375-720 (λ_{ex} 355); 390-720 (λ_{ex} 370) and 430-800 nm (λ_{ex} 410). The slit width for excitation was 2 nm, and it was kept at 15 nm for emission. Concentrations for all amino acids were 0.3 M, whereas, for poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2, it was 5 mg/mL, 7.5 mg/mL and 50 μM, respectively. All

the spectra were normalized to unity for comparison between various proteins and amino acids for various luminescence properties. Moreover, the area under the curve for all the samples was integrated for each spectrum and plotted versus the corresponding exciting wavelength.

Luminescence Emission of switch peptides was carried out with 50 μM of peptides and 3 different excitation wavelengths: 295 nm, 340 nm and 370 nm were used, and emission spectra were collected between 315-570nm; 360-660 nm; 390-720 nm respectively. Slit width: excitation = 2 nm, emission = 15 nm. All readings were taken in duplicates, and 10 scans were recorded for each and averaged later. Moreover, the area under the curve for all the samples is integrated for each spectrum and is plotted versus the corresponding exciting wavelength.

Assembly of the 10 μM HBV Cp dimer in 50 mM HEPES with 0.25 M NaCl and 0.5 M NaCl, pH 7.4 (assembly condition), was monitored by collecting the ProCharTS luminescence for 30 minutes and was compared with dimer with the same concentration in 50 mM HEPES with no salt, pH 7.4 (non-assembly condition). Various excitation wavelengths used were: 340 nm, 355 nm, and 370 nm, and the corresponding luminescence collected from 370-560 nm, 380-570nm, 400-570 nm with excitation slit at 2 nm and emission kept at 15 nm; Tryptophan fluorescence for the same was carried out with excitation at 295 nm and emission collected from 315-450 nm. NATA was used as a control for the Tryptophan fluorescence study. Moreover, the change in fluorescence intensity at 295 nm excitation while keeping the emission fixed at 320 nm was monitored by data acquisition in kinetics mode with an integration time of 0.1s and was allowed to run for 1800 seconds until the reaction was almost over. The slit width was kept at 2 nm and 5nm for excitation and emission, respectively. Also, data acquisition in kinetics mode monitored the change in ProCharTS luminescence intensity at 340 nm excitation while keeping the emission fixed at 420 nm. The spectra were recorded with an integration time of 0.1s for 1800 seconds until the reaction was almost over. The slit width was kept at 2 nm and 15 nm for excitation and emission, respectively. Additionally, the area under the curve was integrated for each spectrum and was plotted versus the corresponding exciting wavelength.

To study the effect of amino acids on the Fluorescence of NATA, NATA in the presence of a different concentration of amino acids: Glycine and Arginine [0-300 mM]; Histidine and Aspartate [0-200 mM] were excited at 295 nm with a slit width of 2 nm and the emission collected in the range of 320-570 nm data with 10 nm of slit width. The concentration of NATA was kept constant at 20 μM in all the samples. Two different integrated areas under the curve

were measured for the integrated Fluorescence, one between 320 nm to 360 nm and another between 320 nm to 570 nm. The former range of Fluorescence was selected since the bandpass filter (340±20 nm) was used for time-resolved measurements. All the samples were freshly prepared in deionized water, and all measurements were made at 23-25 °C in duplicates.

Steady-state Fluorescence for the Dansyl-labelled HBV dimers was measured at 340 nm excitation with a slit width of 2 nm and emission collected between 355-600 nm with a slit width of 10 nm. Dansyl-labelled HBV dimers mixed with 1:15 and 1:30 dilutions of excess unlabelled HBV dimers (total concentration kept at 10 μM), were allowed to aggregate in 50 mM HEPES with 0.5M NaCl, pH 7.4. Similarly, IAEDANS-labelled HBV dimers mixed with 1:10 and 1:15 dilutions of excess unlabelled HBV dimers (total concentration of 10 μM) were allowed to aggregate in the same buffer.

FRET studies of Labelled HBV dimers were carried out by measuring the Tryptophan fluorescence in labelled and unlabelled dimers (1 μM of dimer). The samples were excited at 295 nm (slit width of 2 nm), and emission was collected between 310-570 nm (slit width of 10 nm).

Fluorescence polarization experiments were performed with 10 μM of HBV Cp in 50 mM HEPES (for dimer) plus 0.5M NaCl (for capsids). Excitation was carried out at 295 and 310nm, and emission was kept fixed at 345 nm. For HBV Cp, the slitwidth used for 295nm was 1 nm(excitation) and 15 nm(emission), whereas, for 310 nm, the slits used were 3 nm (excitation) and 15 nm (emission). For HSA and PEST, the slitwidth used for 295nm was 1 nm(excitation) and 15 nm(emission); however, for 310 nm, the slits used were 5 nm (excitation) and 15 nm (emission). 3 independent readings were taken for each measurement with 10 scans each.

Luminescence Linearity: Linearity in luminescence was carried out at λ_{ex} 355 nm with similar concentrations of all samples as used for acquiring emission spectra. The luminescence spectra thus collected were integrated and linearly fitted with varied concentrations using OriginPro 8.5 software.

Luminescence Excitation Spectra: Excitation spectra for Arginine HCl, Arginine, Histidine and Aspartate KCl, poly Arginine HCl, poly Aspartate NaCl, Symfoil PV2, and all the switch peptides were collected at 3 different emission wavelengths: 425, 450 and 480 nm and the excitation wavelength were scanned between 230-410, 240-430 and 260-460 nm respectively.

The excitation slit width of 2 nm and an emission slit width of 5 nm was kept for all the readings.

All steady-state luminescence was measured either using a Fluoromax-4 Spectrofluorometer (Jobin-Yvon Horiba Inc., USA) equipped with motorized polarizers or FluoroMax3 (Jobin Yvon, Horiba). To study HBV capsid assembly and the effect of charged residues on the Fluorescence of NATA, Fluoromax-4 Spectrofluorometer was used, whereas, for the rest of the experiments, FluoroMax3 was used. The luminescence intensity from buffers was subtracted from the emission spectra of the sample. Samples were always freshly dissolved in deionized water or respective buffer, and all the measurements were done in a 500 µL quartz cuvette with 10 mm path length (Hellma; Z802875) at room temperature. Emission spectra for all samples were corrected with a data acquisition set as an S1c/R signal. Scanning was carried out with an integration time of 0.1s and a wavelength increment of 1 nm.

2.3.9. Stokes shift

We calculated the excitation wavelength-dependent Stokes shift (in wavenumbers) by the following equation-

$$Stokes\ Shift(cm^{-1}) = \left(\frac{1}{\lambda_{ex}} - \frac{1}{\lambda_{em}}\right) \qquad (2.49)$$

where λ_{ex} is the excitation wavelength, and λ_{em} is the wavelength of emission maxima, both in cm.

For the emission spectra of Arginine HCl, Arginine, Histidine and Aspartate-KCl, polyArginine-HCl, polyAspartate-NaCl and Symfoil PV2, Stokes shift for the following exciting wavelengths: 280, 295, 310, 340, 355, 370, and 410 nm were measured.

2.3.10. Quantum yield

Quantum yield for Arginine HCl, Arginine, Histidine and Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl and Symfoil PV2 at 355 nm excitation were calculated using the following equation-

$$\phi_f^i = \frac{F^i f_s n_i^2}{F^s f_i n_s^2} \phi_f^s \qquad (2.50)$$

Here, 9, 10 diphenyl anthracene (DPA) was used as a reference[152]. ϕ_f^x, F^x , n_x , f_x denote the luminescence quantum yield, integrated luminescence, refractive indices, and absorption factor, where $f_x = 1 - 10^{-Ax}$ (A = absorbance at 355 nm). Subscript or superscript i and s denote measurements for the sample and the reference, respectively.

All the proteins were prepared in deionized water, while for the 9,10-DPA, cyclohexane (99%) was used. For both sample and reference (DPA), the absorbance values at 355 nm were kept below 0.04. Corrections for the refractive index of cyclohexane (DPA) and aqueous samples were carried out with a refractometer. Excitation was done at 355 nm (slit width of 1 nm), emission was collected between 370-690 nm (slit width of 5 nm), and the integrated luminescence area was calculated. All the measurements were done in triplicates at room temperature.

2.3.11. ANS binding assay

8-Anilinonaphthalene-1-sulfonic acid (ANS) binding measurement was carried out by measuring the steady-state ANS fluorescence intensity. 20 μM of aggregated SwPs were collected at various time intervals and mixed with 20 μM ANS (from the stock) in a 1:1 ratio, and their spectra were recorded. Moreover, a spectrum with 20 μM ANS alone mixed in the same buffer was taken as control. The buffer was kept the same as the phosphate buffer (pH 7.2) used for aggregation kinetics. ANS was prepared by using an extinction coefficient of 4,900 $M^{-1}cm^{-1}$ at 350 nm[153]. Excitation was done at 380 nm, and emission was collected between 400 nm to 700 nm. Slit width for excitation was kept at 2 nm and emission at 5 nm. The background luminescence from the buffer was subtracted from emission spectra for each sample. All readings were taken in duplicates, and 2 spectral scans were recorded for each reading.

2.3.12. ThT assay

For Thioflavin T(ThT) assay, a stock of ThT (~700 μM) was prepared in deionized water using an extinction coefficient of 31600 $M^{-1}cm^{-1}$ at 412 nm[154]. All readings were taken in duplicates,

and 4 scans were recorded for each. 20 μM of aggregated SwPs were collected at various time intervals and mixed with 20 μM ThT (from the stock) in a 1:1 ratio, and the spectra were recorded. Also, a control experiment with only 20 μM ThT alone mixed in the same buffer was measured. The buffer was kept the same as the phosphate buffer (pH 7.2) used for aggregation kinetics. Samples were excited at 450 nm (slit width = 2 nm) and emission is collected from 465-550 nm (slit width = 5 nm).

2.3.13. Time-resolved fluorescence

All the luminescence time-resolved intensity decay measurements were done by time-correlated single-photon counting (TCSPC) instrument, using the Delta-Pro™ equipped with a motorized polarizer, supplied by Horiba Scientific, UK. The emission of all samples was measured at a magic angle position of 54.7° with an emission polarizer. Signal acquisition was performed with a Photo Multiplier Tube (PPD-900; Horiba Scientific). The luminescence intensity decay for all the samples was collected in 4096 channels with 28.38 ps/channel resolution. IRF was collected using a scattering suspension of chalk powder. The absorbance of samples either prepared in deionized water or respective buffers were kept below 0.08 at excited wavelengths to avoid inner filter effects. All lifetime measurements were done at room temperature (25 °C), and measurements were recorded in triplicates. All measurements were done with 10 mm path length quartz cuvette (Hellma; Z802875) at 25 °C.

Fluorescence lifetime of NATA + Symfoil 4P/PV2: All samples were excited with a pulsed light source, 295 nm DeltaDiodeTM with ~810ps FWHM Instrument Response Function (IRF); and 20 MHz repetition rate. Emission was collected using a 340±20 nm bandpass filter (Chroma, ET340/40×). Total peak counts were collected up to 15,000 counts.

To measure the luminescence intensity of switch peptides: All samples were excited with a pulsed light source, 340 nm Delta Diode™ having 800ps FWHM IRF and a repetition rate of 20MHz. Emission was measured using a 370 nm long-pass filter (KV370) to cut the excitation light. Total peak counts were measured up to 10,000 counts.

Fluorescence lifetime of Symfoil 4P/PV2; all amino acids and homo-polypeptides: All samples were excited with both pulsed light source, 295 nm DeltaDiode™ with ~810ps FWHM Instrument Response Function (IRF); and 20 MHz repetition rate and 340 nm Delta Diode™

having 800ps FWHM IRF and a repetition rate of 20MHz. Subsequently, emission was collected using a 320 nm long-pass filter (Schott; WG320) or 370 nm long-pass filter (KV370), respectively. Total peak counts were collected up to 15,000 counts.

2.3.14. Luminescence decay analysis

Extraction of decay parameters for all the luminescence intensity decays was done using discrete analysis software (kindly provided by Dr. N. Periasamy, Tata Institute of Fundamental Research (TIFR), Mumbai, India)[155] as mentioned in **Section 2.1.3.4.1.** For NATA+Symfoil experiments, both 1 and 2 exponential models were used, whereas, for the switch peptide aggregation, these two models could not fit the data, so 3 exponential fit was used. The discrete analysis generated various parameters from the intensity decay: individual lifetime components with their respective amplitudes, their mean lifetime along with the reduced chi-square. Non-linear least-squares analysis[156] was used to extract these parameters (α_i and τ_i) in successive iterations. Also, the Maximum Entropy method was used to supplement the data [138,139] as mentioned in **Section 2.1.3.4.2.** In this approach, the luminescence intensity decay of each sample was fitted for 100 exponentials. The lifetime range 0.01ns to 20ns of with 0.001 as an initial value was used for flat distribution. The goodness of fit was evaluated by analyzing 2 parameters: the reduced chi-square χ^2 closer to 1 and the randomness of residuals.

2.3.15. Circular Dichroism

A CD spectrometer (Make: Jasco, Model J-1500, Jasco Inc., Maryland, USA), was used to obtain CD spectra. Spectra were collected starting from 190 to 260 nm, keeping scan speed at 100 nm/minute; data pitch at 0.1 nm and bandwidth at 2 nm. 2 mm quartz cuvette (Starna Scientific Ltd) was used for measurements. Blank subtraction for deionized water or buffers was corrected before data acquisition. A thermoregulator is equipped with the CD spectrometer for stabilizing temperature at any defined temperature. For each sample, two independent scans were carried out at 25°C.

For SwP aggregation study, 50 μM and 200 μM of peptide concentrations (depending on the individual SwP peptides) in phosphate buffer (pH 7.2) in various time intervals were used for the CD study. All the readings were done in duplicate scans and at room temperature.

For HBV dimer and capsid study, 2 μM of concentrations of dimer were used. CD of the capsids (2 μM) collected in the purification process in the void volume in Tris-HCl (pH 7.4), along with the dimers (2 μM) in HEPES (pH 7.4) and the capsids (1.8 μM) assembled by adding 0.5 M NaCl were performed. For each sample, three independent scans were carried out.

2.3.16. Size Exclusion Chromatography

ÄKTA purifier™ FPLC (GE Healthcare) chromatography system was used for Size exclusion chromatography. HiLoadTM 16/600 SuperdexTM 200pg (GE Healthcare, 28-9893-35) column was used to purify the HBV dimers to check the capsid assembly formation. This column has a resolution of 10-600 kDa, and the bed volume of the column is 120 mL. SEC was basically carried out thrice during HBV purification. First, the protein pellet was dissolved in 100 mM Tris-HCl, pH 7.5, 100 mM NaCl, 50 mM sucrose, and 2 mM DTT (SEC purification buffer) and filtered and passed the column at 1mL/min and the elute in the void volume (~45 min) was collected which contains the HBV Cp 149 capsid. Second, the purified HBV Cp dimer was passed through the SEC column equilibrated in 100 mM sodium bicarbonate, pH 9.5, 2 mM DTT and pure HBV Cp 149 dimers were collected at around 82 min. Blue Dextran (2,000 kDa) was passed through the column, and its elution time (void volume of column) was measured to be ~45.34 mL. Various protein standards used were: alcohol dehydrogenase (150 kDa), bovine albumin serum (66 kDa), carbonic anhydrase (29 kDa), and hen egg-white lysozyme (14.3 kDa). They were also run in the same elution buffer (25 mM Tris.HCl (pH 7.4, 2 mM DTT) and their elution time were collected, and a standard curve was obtained from those protein standards. Subsequently, the column was equilibrated with the sample elution buffer (25 mM Tris.HCl (pH 7.4); 2 mM DTT), and then HBV dimers were loaded, and its molecular weight was estimated from the standard curve.

Third, the aggregation of HBV Cp dimer was carried out in 50 mM HEPES with 0.5 M NaCl (pH 7.4) with 10 μM dimers, and 1 mL was loaded in the column to check aggregation and the peak was collected again at the void volume, i.e., ~45 min thus confirming capsid assembly formation. The column was washed with 1-2 column volume of degassed water between the usage of different buffers. Also, the column was generally equilibrated with 2 column volumes of equilibration buffer. A flow rate of 1 mL/minute was maintained for all samples and standards throughout all individual experiments. Moreover, all the buffers were carefully degassed and filtered with a disc filter attached to a vacuum pump.

2.3.17. Dynamic Light Scattering (DLS)

LiteSizer 500 (Anton Paar) was used to carry out Dynamic Light Scattering measurements. Scattering experiments for the sample were done in backscatter mode with an equilibration time of 10 seconds each. Each measurement was performed with 10 processed runs, each with a time duration of 5 seconds, and the distribution was volume-weighted. All the measurements were done in triplicates and were performed at room temperature (25 °C). Before the measurements, all samples were filtered) to remove dust particles through a 0.22 µm syringe filter (GE healthcare life sciences Whatman TM Cat no.9908-2502) which can otherwise give erroneous results. DLS for dimer was carried out with a dimer concentration of 4 µM, whereas for capsids, 4 µM capsids were used from the capsids assembled in 50 mM HEPES with 0.5 M NaCl (pH 7.4) with 10 µM dimers.

2.3.18. 90° Light scattering

90° light scattering was performed using Fluoromax-4 Spectrofluorometer (Jobin-Yvon Horiba Inc., USA) equipped with motorized polarizers. The scattering kinetics was observed in capsids (10 µM) in 50 mM HEPES with 0.5 M NaCl (pH 7.4) at 320 nm and 340 nm excitation wavelengths, and scattering was collected at 90° at the same wavelengths. The data acquisition was carried out in kinetics mode with an integration time of 0.1s. The assembly was allowed to run for 1800 seconds until the reaction was almost over. The slit width was kept at 1 nm and 5 nm for excitation and emission, respectively.

2.3.19. Labelling of HBV Cp149 with Dansyl Chloride

Labelling of HBV Cp 149 with Dansyl chloride was carried out as described previously[157]. Briefly, the labelling reaction was started by adding 30 µL of 14 mM stock of Dansyl chloride prepared in dimethyl formamide (DMF) to 0.67 mL of 40 µM of HBV Cp 149 dimer prepared in 0.1 M bicarbonate buffer (pH 9.3). This reaction continued for 3 hours at 4 °C with continuous steering. After incubation, the labelled sample was passed through PD-10 columns (GE Healthcare) pre-equilibrated with 50 mM HEPES buffer, pH 7.5 and elution was performed where the unlabelled dyes separated. Protein and labelled dye thus collected were quantified by measuring the absorbance values at 280 nm (for dimer protein) and 339 nm (for dye) using a double beam Lambda-25 UV-Vis spectrophotometer (Perkin Elmer, USA). The

extinction coefficient of HBV Cp 149 is 60,900 $M^{-1}cm^{-1}$ at 280 nm[124], while for conjugated dye, it is 3370 $M^{-1}cm^{-1}$ at 339 nm[158,159]. The concentration of dye and protein was measured spectrophotometrically using a 10 mm path length quartz cuvette (Hellma; Z600210) at 25 °C. The dye to dimer protein (D/P) ratio was around 1.4.

2.3.20. Labelling of HBV Cp 149 with 1,5-IAEDANS

Labelling of HBV Cp 149 with 1,5-IAEDANS (1,5-IAEDANS, 5- ((((2-iodoacetyl) amino) ethyl) amino) naphthalene-1-sulphonic acid) was carried out as recommended by Molecular Probe with minor modifications[160]. Similar to that of Dansyl chloride labelling, the labelling reaction was started by adding 14 µL of 16 mM stock of Dansyl chloride prepared in dimethyl formamide (DMF) to 0.48 mL of 40 µM of HBV Cp 149 dimer prepared in 25 mM Tris. buffer (pH 9.3). The duration of the reaction and elution buffers were the same as Dansyl chloride labelling. Protein and labelled-dye concentrations were calculated by measuring the absorbance values at 280 nm (for dimer protein) and 337 nm (for dye). The extinction coefficient of HBV Cp 149 is 60,900 $M^{-1}cm^{-1}$ at 280 nm[124], while for conjugated dye, it is is 6100 $M^{-1}cm^{-1}$ at 337 nm[161]. The D/P ratio was calculated to be around 0.7.

2.3.21. Steady-state fluorescence anisotropy

Anisotropy measurements were performed using Fluoromax-4 Spectrofluorometer (Jobin-Yvon Horiba Inc., USA) equipped with motorized polarizers. For measurements, three different conformations of HBV dimers: dimers in the denatured state using 6 M Gdn.HCl, dimers in native form and dimers in capsid state were explored. For all tryptophan steady-state fluorescence anisotropy measurements of HBV Cp149, samples were excited at fixed 295 nm (slit width = 2 nm) and emission was collected at fixed 323 nm (slit width = 10 nm). For collecting steady-state anisotropy of IAEDANS labelled with cysteine of HBV Cp, all samples were excited at fixed 340 nm (slit width = 2 nm), and emission was collected at fixed 483 nm (slit width = 10 nm). For collecting steady-state anisotropy of Dansyl labelled with lysine of HBV Cp, all samples were excited at fixed 340 nm (slit width = 2 nm), and emission was collected at fixed 437 nm (slit width = 10 nm). Anisotropy measurements were repeated at least three times each with an average of 10 independent measurements. All experiments were performed at room temperature (25 °C) in a 700 µL quartz cuvette with a 10 mm pathlength [Hellma (Z802875)].

2.3.22. Transmission Electron Microscopy (TEM)

Freshly assembled HBV Cp 149 capsids were adsorbed to freshly glow-discharged grids on a 300-mesh copper substrate. Excess samples were removed by placing wet tissue paper at one of the corners of the grids and washed with water immediately. Next, the samples were stained with freshly prepared 2% uranyl acetate salt[162] and subsequently washed with water to remove the excess salts. Samples were then kept overnight. Samples were visualized on a Transmission Electron Microscope (TEM), Make JEOL, Model: JEM 2100 retrofitted with a charge-coupled device camera. Images captured have dimensions of 500 nm x 500 nm.

2.3.23. Atomic Force Microscopy (AFM)

Freshly assembled HBV Cp 149 capsids were placed on a glass slide cleaned with lint-free tissue paper and allowed to stay overnight. The glass slides were then stuck with double-sided carbon tapes and placed under the Atomic Force Microscope (Make: Oxford Instruments, Model: Cypher S). The samples under the AFM and the force between a probe and the sample were measured in non-contact mode. Images captured have dimensions of 2 μm x 2 μm.

Chapter 3

ProCharTS arising from Arginine, Aspartate amino acids observed in peptides & proteins which are devoid of Lysine residues

3.1. Introduction

Spectroscopic techniques, like absorption or fluorescence, use various extrinsic (fluorescent dyes) and intrinsic chromophores (aromatic amino acids, prosthetic groups, peptide bonds) to monitor proteins' structure, function, and dynamics. Since extrinsic probes can alter protein structure, intrinsic chromophores like Tryptophan (Trp) are often used to study protein structure-function relationships. However, unusual absorbance and even luminescence have been reported to arise from non-aromatic amino acids in proteins. Multiple hypotheses for the origin of such novel absorbance and fluorescence have been postulated, but the reason remains unclear. Recently, a study on the intrinsic UV/Visible absorption was performed in an extensively charged protein α_3C with no aromatic amino acids[23]. Photoinduced electron transfer from various Lys-Glu, Lys-Lys, and Glu-Glu contacts present in the protein was thought to be the reason for the intrinsic absorption in the range 250-800 nm and was termed as Protein Charge Transfer Spectra (ProCharTS). Intrinsic luminescence due to ProCharTS was also observed in monomeric proteins rich in Lysine and Glutamate residues. Moreover, ProCharTS has been used as a tool to monitor the structural changes in Intrinsic Disordered proteins and the aggregation of Human Egg White Lysozyme (HEWL)[43].

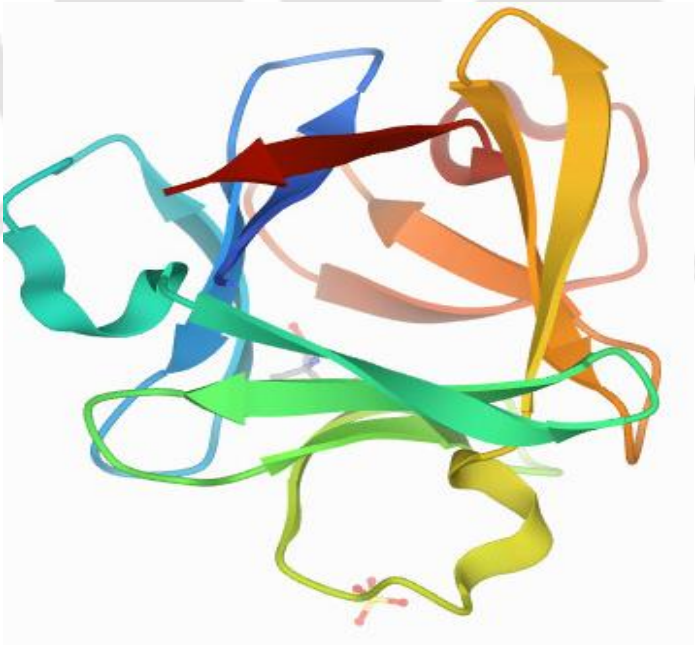


Figure 3.1: PDB structure of Symfoil 4P/PV2. (PDB ID: 4D8H)

The majority of the proteins studied so far involve intrinsic ProCharTS absorbance/luminescence arising predominantly from lysine and glutamate amino acid residues. Apart from the initial works by Swaminathan and group in 2017 on α_3C and the other proteins, time-dependent density functional theory (TDDFT) calculations were performed on various Arginine, Aspartate, and Histidine side chains, as well as on phosphorylated Serine and Threonine residues. However, no reports of *in vitro* studies of ProCharTS observed in proteins/peptides rich in other charged amino acids like Asp, His, and Arg while lacking Lys and Glu have been reported. In this study, we report ProCharTS absorbance and luminescence in Arginine, Aspartate, and Histidine-rich protein/peptides for the first time. A highly charged protein rich in Arg/Asp with no lysine and any aromatic residues was selected, consisting of 126 amino acids with 30% charged content, the structure shown in **Figure 3.1**. Moreover, the novel absorbance and luminescence arising from these amino acids and their respective homo-polypeptides were investigated to understand the phenomenon better. Subsequently, we tried to find the correlation among the observed optical properties in all these amino acids, homo-polypeptides, and Symfoil 4P/PV2. Also, whether the observed absorbance and luminescence arising from these amino acids mimic the optical features of Lysine and Glutamate were analyzed to decipher the origin of such a novel intrinsic phenomenon. It is to be noted that Symfoil 4P/PV2 has been represented as Symfoil PV2 throughout in the thesis.

3.2. Results and Discussions

3.2.1. Symfoil PV2 expression and purification

In this study, only Symfoil PV2 protein has been purified, while the other peptides and amino acids have been purchased. So, the expression and purification of Symfoil PV2 have been carried out and further analyzed. Symfoil PV2 has been expressed in *E. coli* BL21 (DE3) and subsequently purified as per the procedure mentioned in *Chapter 2*. **Figure 3.2A** shows the expression of Symfoil PV2 in 15% SDS PAGE. It could be seen that the protein got expressed (a distinct band near 14 kDa) in all induced conditions whereas, in the uninduced control, no such bands were visible. The result suggested the protein has been expressed successfully under induced conditions with 0.1-1.0 M IPTG concentrations. Consequently, the purity of the protein is shown in **Figure 3.2B.** The single band for the purified Symfoil PV2 is clearly visible in lane 6, showing a molecular weight of around 14.3 kDa.

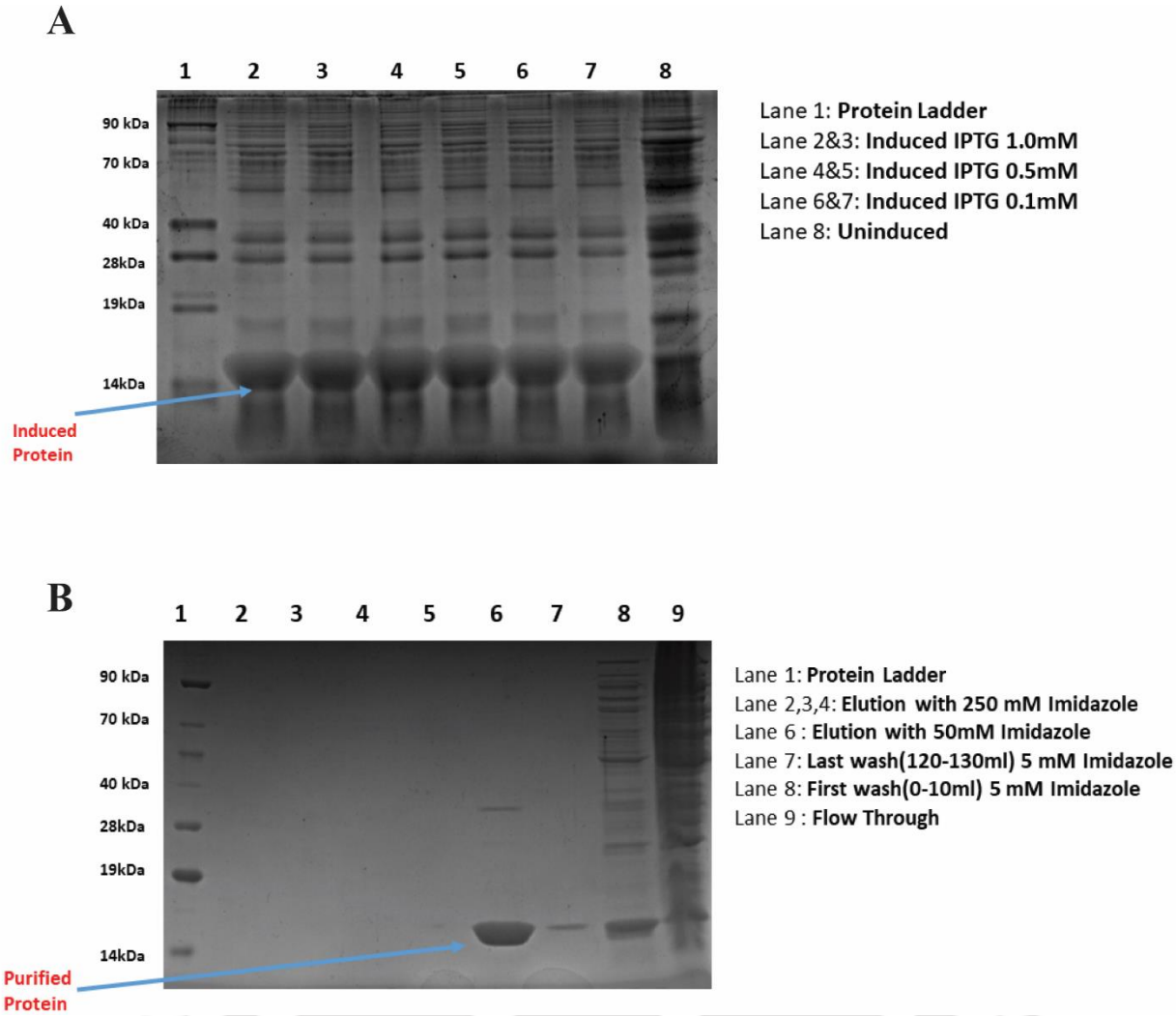


Figure 3.2: Expression (A) and purification (B) of Symfoil PV2 shown in 15% SDS PAGE.

3.2.2. MALDI of Symfoil PV2 and homo-polypeptides

To further cross-check the purity of the protein, MALDI was performed. In **Figure 3.3C**, a clear spike with an m/z value of 14451.13 Da (z = 1) and 7223.847 Da (z = 2) could be seen, suggesting the experimental molecular weight to be 14451.13 Da, which is quite close to the theoretical mass of 14304.58 Da. The additional mass of 150 Da may be due to some charge ions or adducts attached in the gaseous phase, like protons or water molecules. Moreover, the MALDI for the Arginine and Aspartate homo-polypeptides, as seen in **Figures 3.3A** and **3.3B**, reflected the typical distribution of m/z peaks for homo-polypeptides, with values ranging in the 5000-15000 Da and 2000-11000 Da, respectively. The weighted-molecular mass was calculated to be 8753 Da and 5714 Da for Arginine and Aspartate polypeptides, respectively.

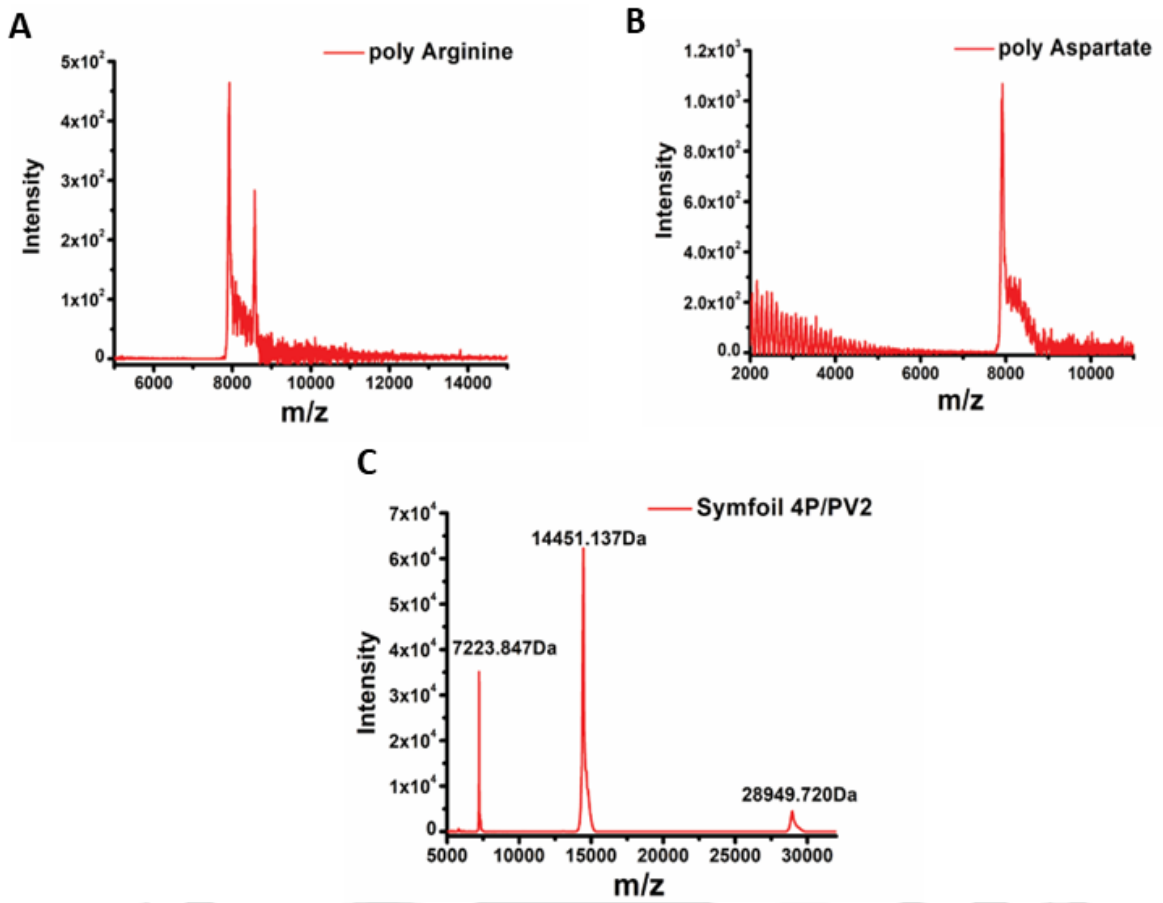


Figure 3.3. MALDI image of (A) poly-Arginine, (B) poly-Aspartate, and (C) Symfoil PV2.

3.2.3. ProCharTS absorbance

ProCharTS absorbance of the amino acids: Arginine Salt, Arginine, and Aspartate Salt were measured for 10 concentrations from 0.1 M to 1.0 M, except Histidine, where 6 concentrations were used, and the dependence of ProCharTS on concentration at different wavelengths were measured, as shown in **Figure 3.4.** Similarly, the absorbance profile of poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 was also recorded, as shown in **Figure 3.5.** 5 different concentrations of poly-Arginine and poly-Aspartate were used to measure ProCharTS. In contrast, for Symfoil PV2, 6 different concentrations between 15-125 μM were taken. All the amino acids, homo-polypeptides, and Symfoil PV2 displayed a linear fit with distinctive slopes having an R^2 value of > 0.97, > 0.98, and > 0.99, respectively. Symfoil PV2 and the homo-polypeptides showed significantly higher slopes than amino acids for any given wavelength, as seen in **Figures 3.4 and 3.5.** Subsequently, the extinction coefficients "ε" were calculated as ProCharTS absorbance per unit increase in concentration is indicative of ε.

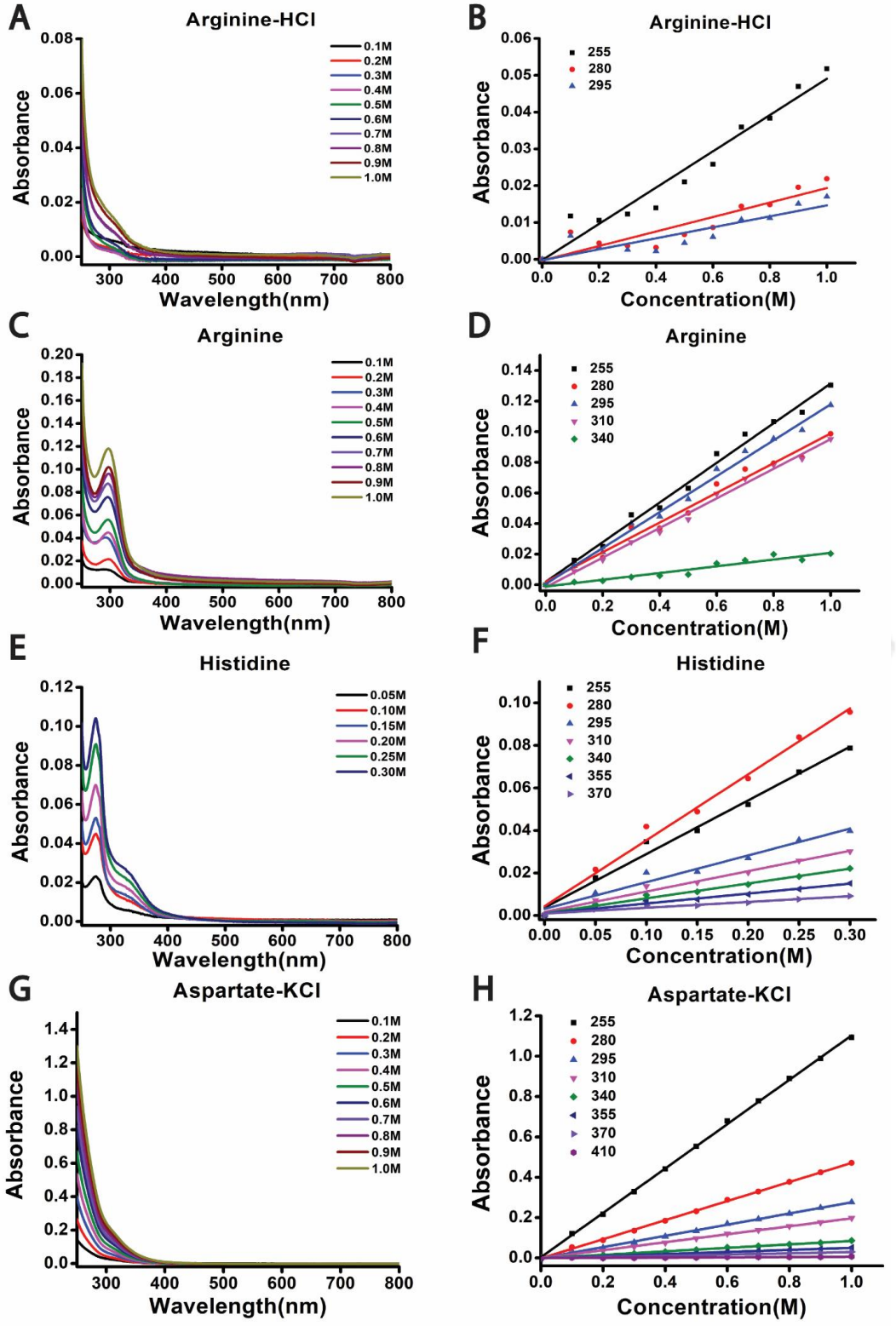


Figure 3.4. The absorbance of Arginine-HCl, Arginine, Histidine, and Aspartate-KCl is shown here.10 different concentrations were used for all samples, except Histidine, which is shown in subfigures A, C, E, and G. The corresponding absorbance versus concentration plots at different wavelengths are shown in subfigures B, D, F, and H.

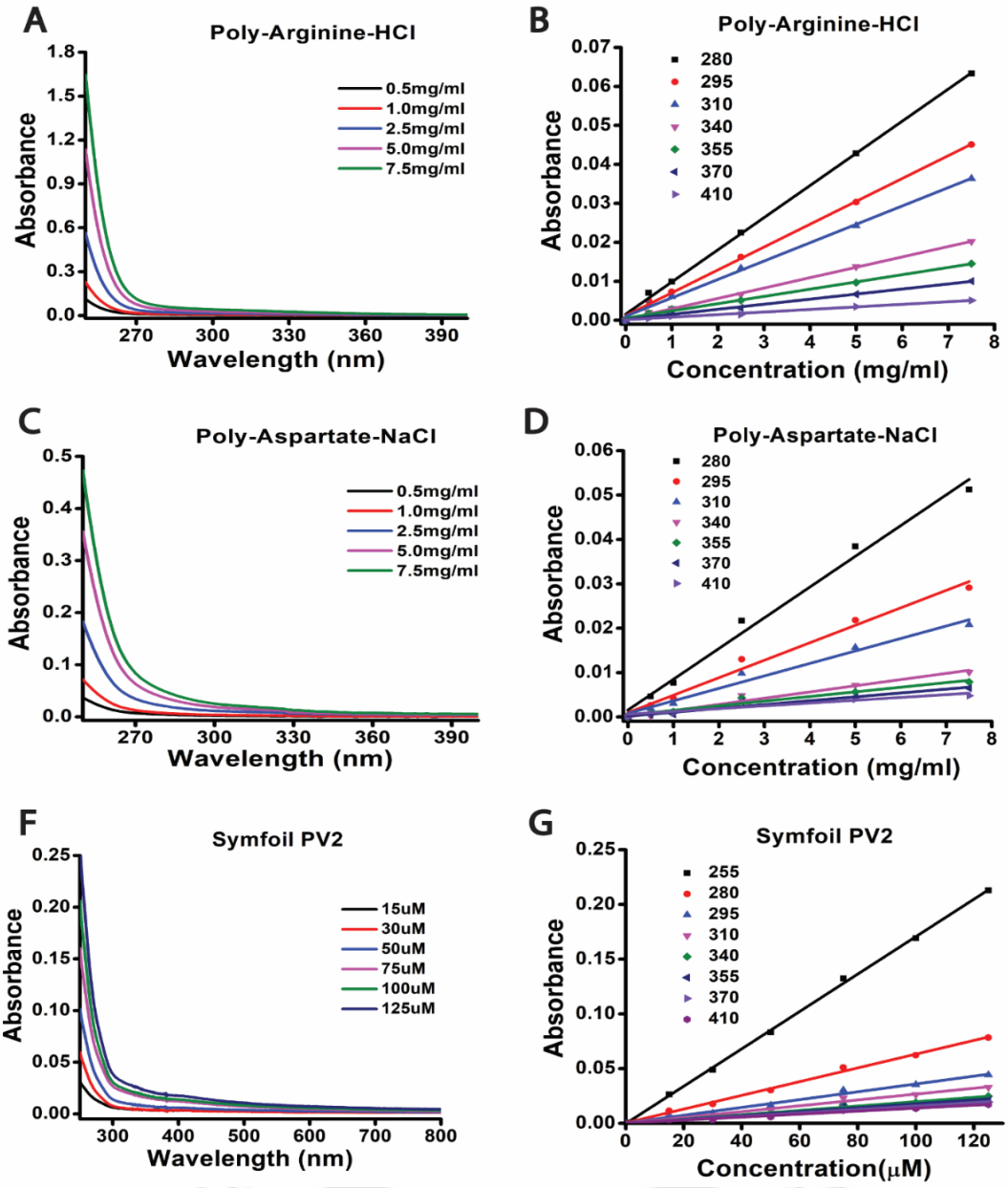


Figure 3.5. The absorbance of poly-Arginine-HCl and poly-Aspartate-NaCl, Symfoil PV2 is shown here. 5 different concentrations were used for polypeptides, and their absorbance versus concentration plot is shown at 7 different wavelengths in subfigures A, C, and B, D, respectively. The absorbance of Symfoil PV2 for wavelengths ranging from 250-800 nm for 6 different concentrations and its absorbance versus concentration for 8 different wavelengths is shown in subfigures F and G.

ProCharTS spectra across all samples displayed a similar profile with decreasing absorbance with increasing wavelength and an absorbance tail till 800 nm. However, Histidine and Arginine showed unusual humps at 280 nm and 300 nm, respectively, the reason for which is not yet known. Also, the ProCharTS intensity was found to vary linearly with concentration suggesting all the samples were in a monomeric state.

Wavelength (nm)	255	280	295	310	340	355	370	410	500	600	700	800
Arginine HCl	0.046 (0.006)	0.017 (0.005)	0.012 (0.004)	0.01 (0.007)	0.003 (0.005)	0.001 (0.003)	-----	-----	------	------	-----	-----
Arginine	0.131 (0.007)	0.098 (0.008)	0.116 (0.007)	0.092 (0.007)	0.018 (0.005)	0.013 (0.004)	0.009 (0.004)	0.004 (0.004)	-------	------	------	-----
Histidine	0.264 (0.004)	0.325 (0.007)	0.137 (0.004)	0.102 (0.001)	0.073 (0.000)	0.050 (0.000)	0.030 (0.000)	0.011 (0.000)	--------	-------	------	------
Aspartate KCl	1.103 (0.014)	0.464 (0.012)	0.269 (0.011)	0.191 (0.008)	0.079 (0.007)	0.045 (0.007)	0.025 (0.006)	0.002 (0.004)	-------	--------	------	------
Poly-Arginine-HCl	1023 (9.3)	88 (21.0)	63 (14.5)	51 (11.5)	27 (5.3)	21.6 (6.4)	13.8 (3.2)	7.3 (2.1)	4.7 (2.6)	3.1 (1.9)	3.2 (2.5)	3.3 (2.9)
Poly-Aspartate-NaCl	264.2 (17.4)	45.9 (5.5)	26.2 (3.5)	19.2 (3.0)	7 (3)	7.7 (1.8)	5.4 (1.6)	6.1 (2.3)	6.0 (3.6)	7.9 (6.8)	10 (11)	7.3 (6.2)
Symfoil PV2	1979 (130)	838 (77)	512 (74)	396 (74)	274 (73)	261 (66)	247 (73)	208 (50)	145 (50)	108 (45)	89 (44)	69 (34)

Table 3.1. The extinction coefficient ($M^{-1}cm^{-1}$) of Arginine-HCl, Arginine, Histidine, Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 is tabulated at various discrete wavelengths. The blank boxes indicate the extinction coefficient could not be calculated precisely due to the very low S/N ratio. The data in parenthesis represents standard deviation.

The molar extinction coefficients were calculated from the slopes of the absorbance versus concentration plots. The molar extinction coefficients for Arginine-HCl, Arginine, Histidine and Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 were calculated as shown in **Figure 3.6**. As can be seen, the amino acids showed relatively lower values of "ε" than the homo-polypeptides and Symfoil PV2, not only at 255 nm but for the entire spectrum. The extinction coefficient for all amino acids, polypeptides, and Symfoil PV2 at multiple wavelengths between 255-410 nm was shown in **Table 3.1**. It could be observed that ε at 280nm (ε_{280}) of all amino acids were far less (<1.0) followed by homo-polypeptides (~ 46-88), while Symfoil PV2 displayed the highest value (~ 838). For some amino acids like Arginine HCl, the ε could not be estimated accurately after 355 nm. In contrast, the values of poly-Arginine-HCl and poly-Aspartate-NaCl were barely measurable at longer wavelengths. Symfoil PV2 showed significantly higher absorption, which extended way till 800 nm. Among

the amino acids, Aspartate-KCl showed the most ProCharTS absorbance, followed by Histidine and Arginine, and the least was calculated for Arginine-HCl.

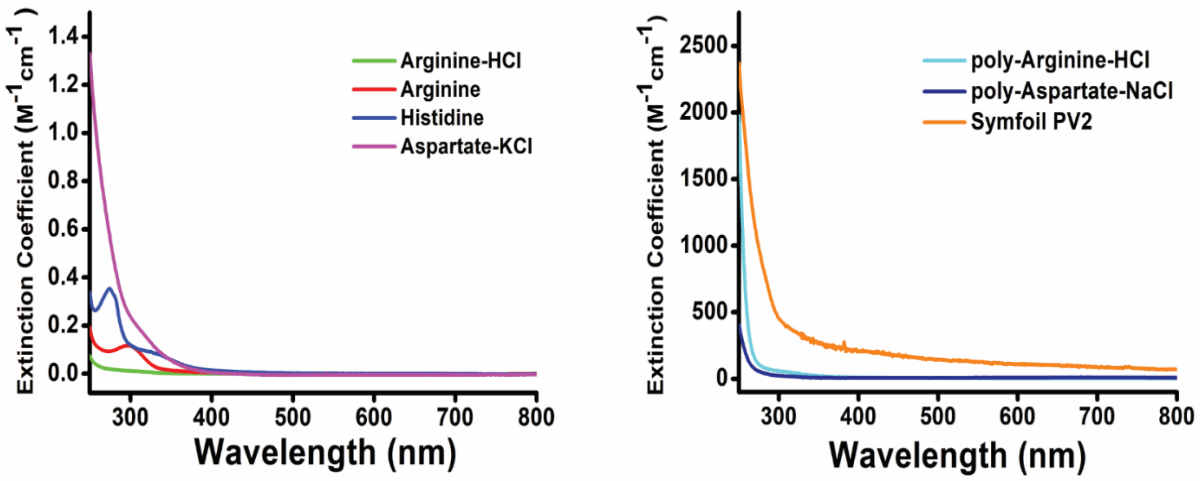


Figure 3.6. Absorption spectra and molar extinction coefficient of different amino acids, homo-polypeptides, and Symfoil PV2 are shown here. The extinction coefficient for amino acids is shown in the left panel, whereas homo-polypeptides and Symfoil PV2 are shown in the right panel.

Free amino acids showed lower ProCharTS due to limited interactions of charged terminals in the 3D space. Even in higher concentrations where they may form clusters, the absence of any peptide bond probably might keep the charged terminals of the amino acids at distances greater than 5-10 Å. However, the polypeptide backbone in the folded structure of Symfoil PV2 and the unfolded homo-polypeptides structures can facilitate charged terminals' interactions and increase the number of contacts among charged residues increasing ProCharTS absorbance.

The long ProCharTS absorbance tail in Symfoil PV2 suggests that charged residue contacts without Lys involvement can also give rise to the characteristic absorbance tail in ProCharTS absorption. A similar absorbance tail in the ProCharTS absorption spectrum was reported for highly charged Lysine-rich monomeric protein α_3C[23] and in some ordered proteins and IDPs[43]. Moreover, aged solution of glycine, where increased interaction of the charged terminal with time, has been reported to show higher ProCharTS absorbance as compared to fresh glycine[163]. Thus, it can be said that ProCharTS absorbance is not only dominant in Lysine-rich proteins but also can be observed in any charged proteins, irrespective of the type of charged amino acids present, even without Lysine residues.

3.2.4. Steady-state luminescence emission

An in-depth study on the luminescence properties of Arginine, Aspartate, Histidine residues in amino acids, homo-polypeptides, and Symfoil PV2 is carried out by both steady-state as well as time-resolved studies. In this section, steady-state luminescence is discussed.

3.2.4.1. Dependence of luminescence on concentration

Luminescence emission spectra for Arginine-HCl, Arginine, Histidine, and Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl and Symfoil PV2 were collected in the range 300-540, 315-570, 330-600, 360-660, 375-720, and 430-800 nm, respectively when excited at the corresponding wavelengths: 280, 295, 310, 340, 355, 370 and 410 nm, as seen in **Figure 3.7.** We observed that although the luminescence at 280 nm excitation showed varied spectra for different samples, the luminescence spectrum for all the remaining excitations behaved similarly. The luminescence intensity stayed consistent till 355 nm and then started declining afterward. It thus suggested that the origin of such novel luminescence from all the samples follows a similar mechanism.

The sensitivity of luminescence to sample concentration was further calculated for all amino acids, peptides, and proteins. Since the emission intensity stayed almost constant till excitation of 355 nm, and drastically drops thereafter, we selected a common excitation wavelength of 355 nm for this study with various concentrations of each sample. The area under the curve from each emission spectrum was calculated, and the integrated luminescence intensity was plotted against concentration, as shown in **Figure 3.8.** All the samples displayed a linear fit with distinctive slopes and an R^2 values > 0.98. Symfoil PV2 showed a maximum slope, followed by poly-Arginine-HCl and poly-Aspartate-NaCl, the values of which were 8.3 x 10^6 (intensity/μM), 4.3 x 10^6 (intensity/μM), and 1.3 x 10^6 (intensity/μM), respectively, whereas the amino acids showed the least slopes. Among the amino acids, Histidine showed maximum slope, followed by Aspartate-KCl, Arginine, and least for Arginine-HCl, the values of which were 5.9 x 10^3 (intensity/μM), 5.7 x 10^3 (intensity/μM), 1.5 x 10^3 (intensity/μM) and 0.44 x 10^3 (intensity/μM), respectively. The luminescence sensitivity followed the trend observed for molar extinction coefficients among the amino acids, homo-polypeptides, and protein. One striking feature was the higher slopes (intensity/μM) observed in proteins and peptides, where the charged terminals are expected to come in close proximity than amino acids, which

indicates ProCharTS luminescence is also reflective of the proximal charged residues. The correlation between both the ProCharTS luminescence and absorbance is described in the subsequent section.

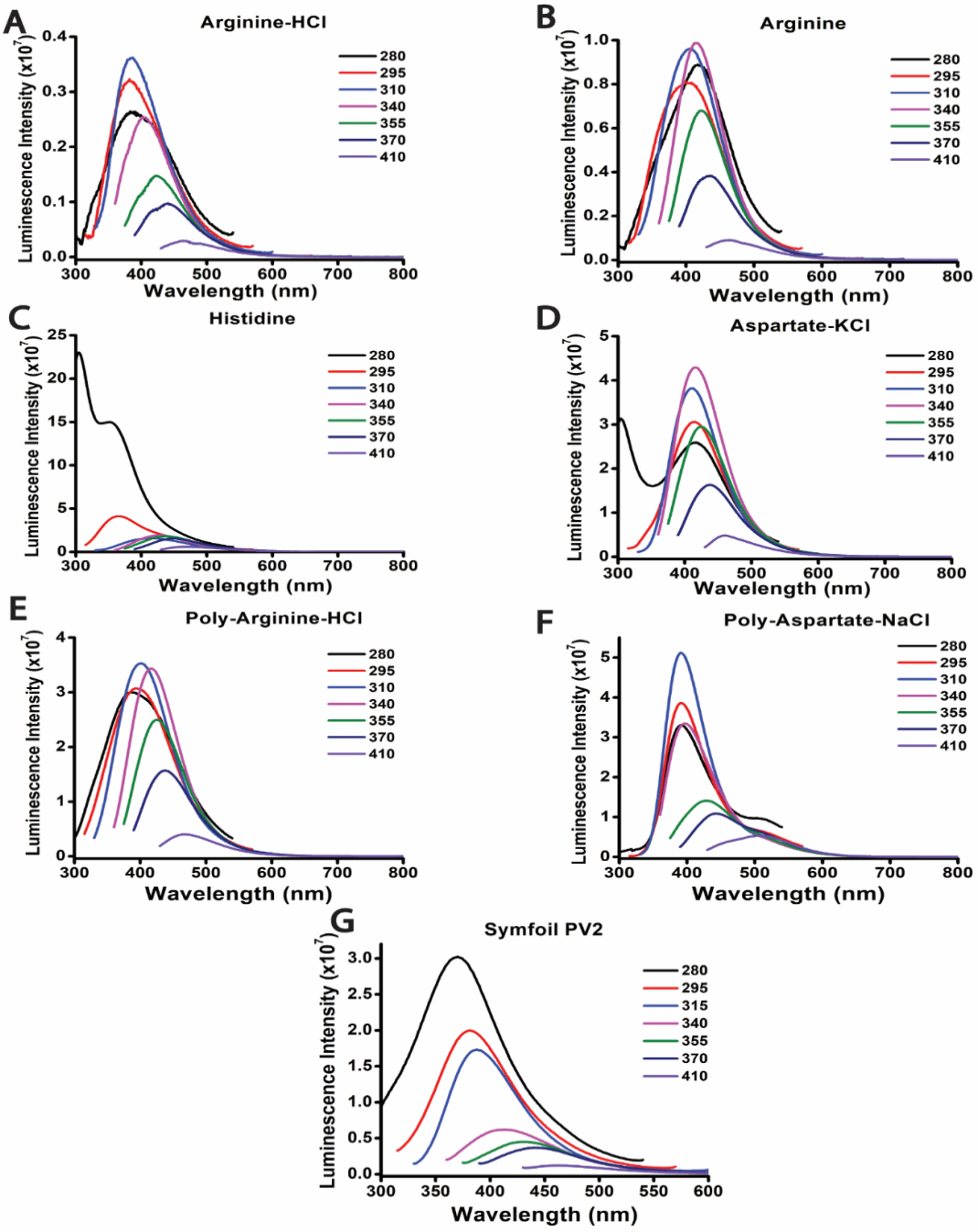


Figure 3.7. The luminescence intensity of Arginine-HCl (A), Arginine (B), Histidine (C) and Aspartate-KCl (D), poly-Arginine-HCl (E), poly-Aspartate-NaCl (F), Symfoil PV2 (G) is shown here. Excited wavelengths were: 280, 295, 310, 340, 355, 370, and 410 nm and their corresponding emission range were 300-540 nm, 315-570 nm, 330-600 nm, 360-660 nm, 375-720 nm, and 430-800 nm.

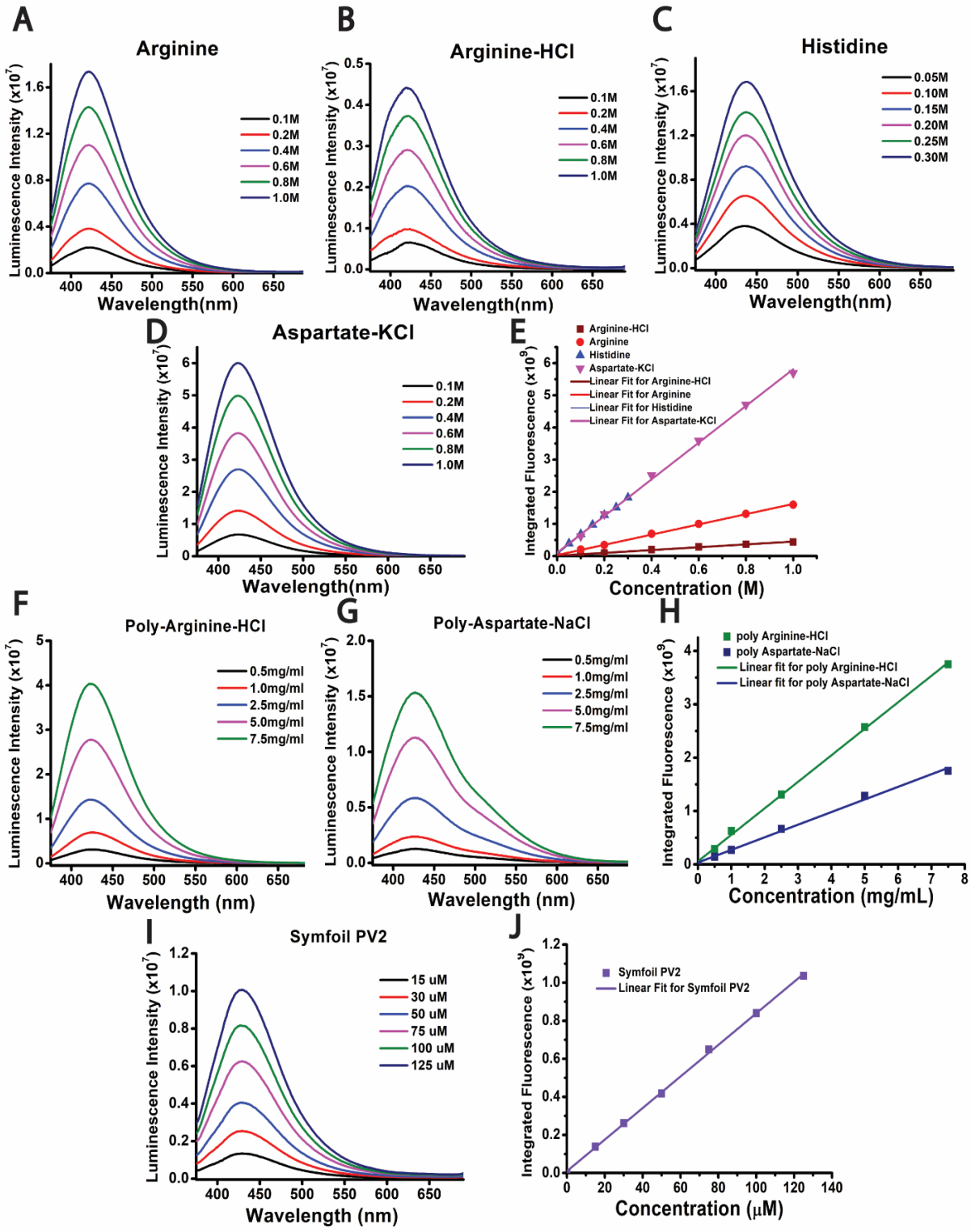


Figure 3.8. Luminescence intensity at excitation 355 nm for various concentrations of Arginine-HCl (A), Arginine (B), Histidine (C), Aspartate-KCl (D), poly-Arginine-HCl (F), poly-Aspartate-NaCl (G), and Symfoil PV2 (I) is shown here. The linearity of the integrated luminescence versus concentration for amino acids, polypeptides, and Symfoil PV2, is shown in subfigures E, H, and J, respectively.

3.2.4.2. Luminescence emission parameters as a function of excitation wavelength

The luminescence emission features were extracted by normalizing individual spectra from the luminescence spectra shown in **Figure 3.7**. The normalized luminescence emission spectra for all the samples at various excitation wavelengths are shown in **Figures 3.9** and **Figure 3.10.**

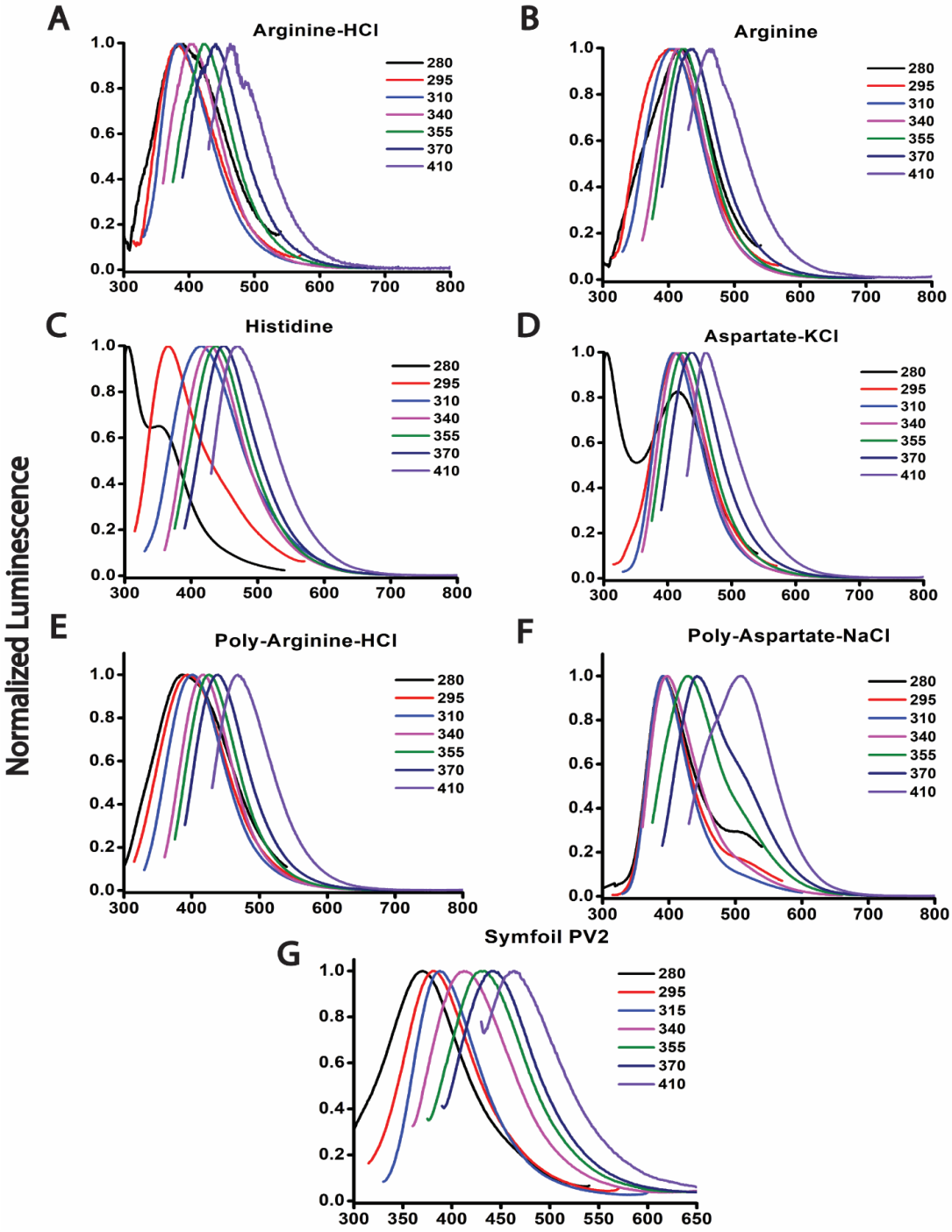


Figure 3.9. Peak intensity normalized luminescence spectra of Arginine-HCl (A), Arginine (B), Histidine (C) and Aspartate-KCl (D), poly-Arginine-HCl (E), poly-Aspartate-NaCl (F), Symfoil PV2 (G) are shown here.

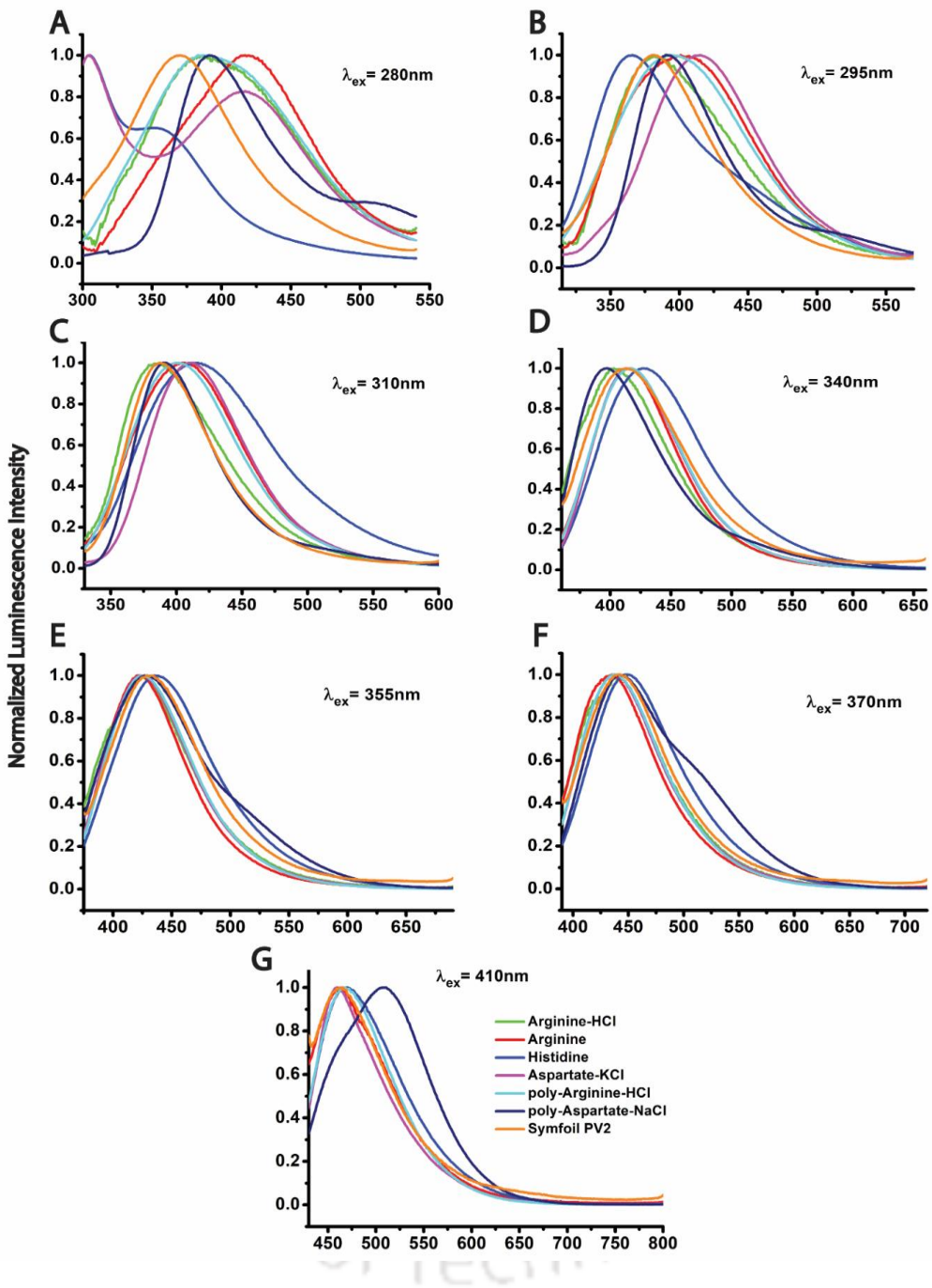


Figure 3.10. Peak intensity normalized luminescence spectra for all samples at different excitation wavelengths: 280, 295, 310, 340, 355, 370, and 410 nm in subfigures A to G, respectively, as a function of excitation wavelengths.

We observed that the luminescence emission spectra for excitation at 280 nm showed poor overlap, but with gradually increasing excitation wavelengths, i.e., for 310, 340, 355, and 370 nm, the individual spectra converged significantly. As already shown in Figure 3.7, the emission intensity decreased with increasing excitation wavelength. The integrated

luminescence area was calculated as a function of excitation wavelength for each sample, as shown in **Figure 3.11A.** It was observed that the total luminescence intensity between 280-340 nm remained nearly constant, but after 340 nm, a distinct decline in emission was visible. We, however, found a strikingly similar pattern of decreasing integrated luminescence in all of the samples. Interestingly, ProCharTS absorbance too showed a considerable decrease after 310-340 nm. This similarity might be because both the ProCharTS absorbance and luminescence underlie the same principle of the electronic charge transfer phenomenon at work.

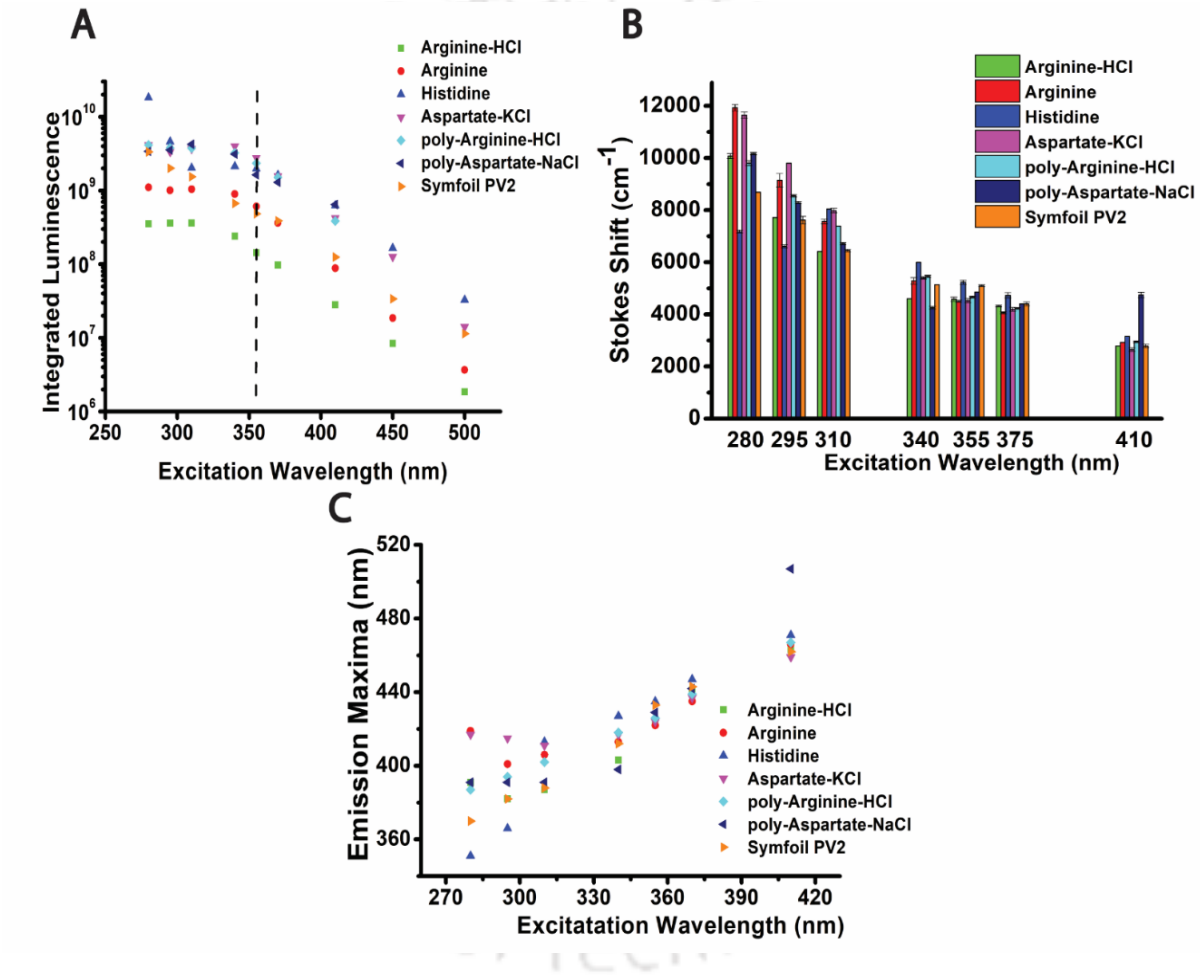


Figure 3.11. Integrated luminescence intensity vs. excitation wavelength (A); Stokes Shift vs. excitation wavelength (B); Emission Maxima as a function of excitation wavelength (C) of Arginine-HCl, Arginine, Histidine and Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 is shown.

The Stokes shift from the emission spectra was calculated for each sample and plotted as a function of excitation wavelength, as shown in **Figure 3.11B**. All the samples showed significantly higher Stokes shift at 280 nm (7000-12000 cm^{-1}) than any known chromophores like the indole ring of Tryptophan, whose Stokes shift is ~ 7143 cm^{-1}. A possible explanation

for this high Stokes Shift is a significantly greater loss in excitation energy, in the form of rapid internal conversion and solvent relaxation, compared to Tryptophan. However, a general decreasing trend of the Stokes shift with increasing excitation wavelength was observed. The Stokes shift decreased steeply from around ~12000 to ~5000 cm^{-1} for 280 and 310 nm. However, from 340 to 370 nm, the decrease is marginal and is somewhat constant (~ 4500 cm^{-1}). At an even higher excitation wavelength of 410 nm, it decreased further to ~3000 cm^{-1}. Strikingly, the changes are consistent across all the samples for each excitation wavelength with the odd outlier.

Although the integrated luminescence decreased with increasing excitation wavelength, when their corresponding emission maxima were plotted with excitation wavelength, a linear trend of emission maxima converging to a narrower spread was evident, as shown in **Figure 3.11C**. Emission spectra were always red-shifted in wavelength irrespective of the excitation wavelength, which is much against the very essence of Kasha's rule. Intramolecular charge transfer was proposed to be the reason behind such observation, as a similar phenomenon has been observed in various other reports[164,165]. Such an increase in the emission maxima with increasing excitation wavelength across all the samples in this study has been consistent with other reports[30,32,45,166]. This excitation wavelength-dependent ProCharTS luminescence suggests the possibility of different singlet CT energy levels from where the electrons after excitation (due to ProCharTS absorbance) can come down to the ground states. A similar observation has also been reported in ligand-fabricated Si nanoparticles[167].

3.2.4.3. Excitation spectra

To get an idea about the nature of the chromophores and understand the excitation bands responsible for the novel ProCharTS luminescence, the excitation spectra of each of the samples were collected, keeping emission wavelength fixed at 3 definite wavelengths: 425, 450, and 480 nm. The choice for these emission wavelengths is simply because most of the luminescence was distributed in the 400-500 nm region when excited at various wavelengths, as seen in **Figure 3.9.** The excitation spectra for all amino acids, peptides, and Symfoil PV2 displayed multiple excitation peaks with broad excitation bands, which showed significant overlap, as shown in **Figure 3.12**. Most of the excitation bands were between 250-400 nm. Both Symfoil PV2 and Histidine showed 2 or 3 excitation bands, whereas Arginine-HCl, Arginine, Aspartate-KCl, poly-Arginine-HCl, and poly-Aspartate-NaCl showed 2 excitation

bands. Broadly, two major excitation bands were evident: one peak at λ <300 nm and the other at λ >300 nm. Our result is in contrast to the previously reported excitation spectra for Lys, Glu amino acids, and Lys/Glu rich proteins, where only one excitation peak was reported[30]. The observance of such multiple excitation bands suggests that the heterogeneous nature of the chromophore is a characteristic of the ProCharTS luminescence. Moreover, the origin of the multiple excitation bands across all the wavelengths appeared similar. It further justified that the heterogeneity in chromophores is conserved across all the amino acids and proteins.

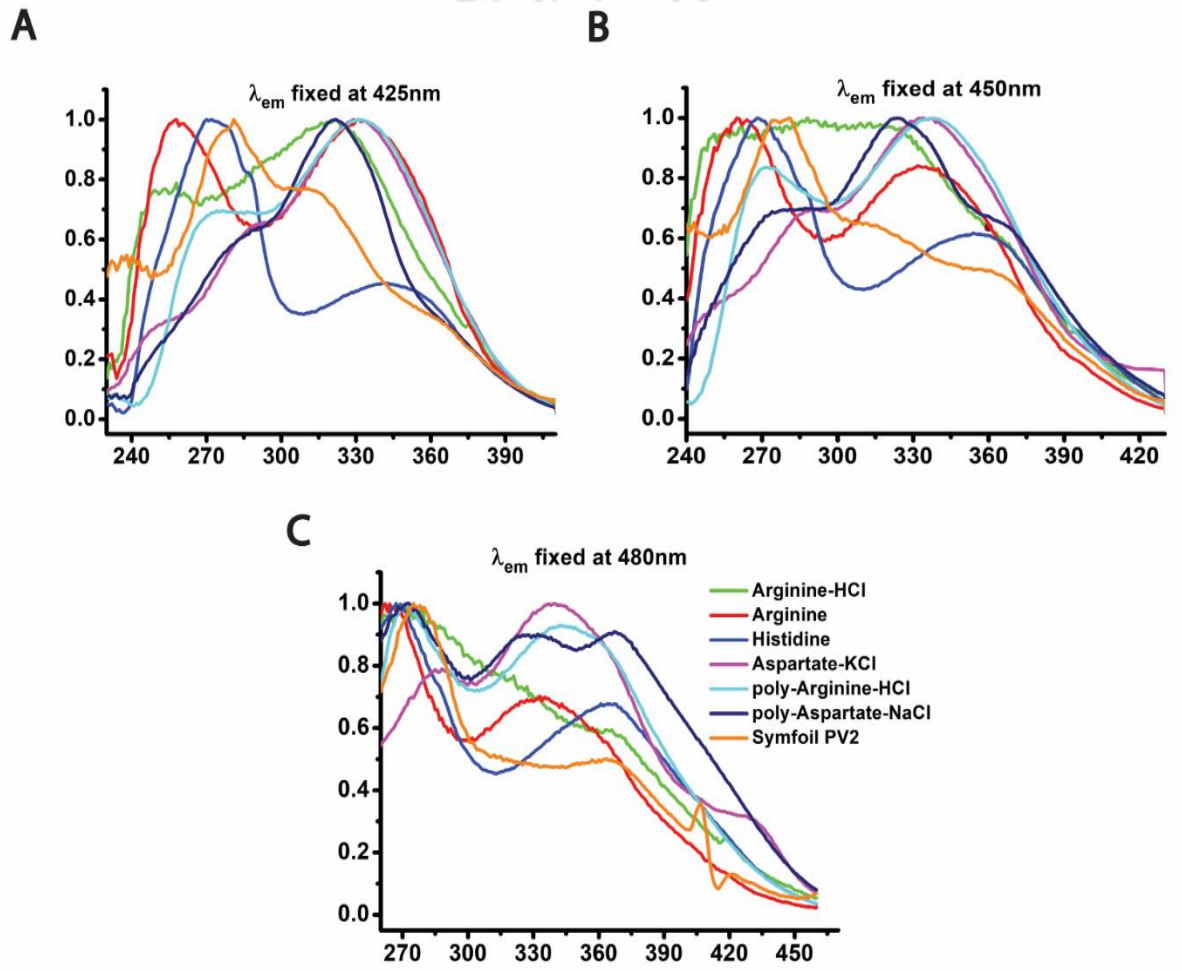


Figure 3.12. Luminescence Excitation Spectra of Arginine-HCl, Arginine, Histidine and Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 for emission wavelengths: 425, 450, and 480 nm in subfigures A, B, and C, respectively.

3.2.4.4. Quantum yield and its correlation with the extinction coefficient

The quantum yield at 355nm excitation (Φ_{355}) for all the samples was calculated using 9,10-diphenyl anthracene (DPA) as reference[152] and is shown in **Figure 3.13A**. It was observed that

the quantum yield values were comparatively lower than those of conventional chromophores like NATA[2] which are around ~0.14. Homo-polypeptides showed the maximum quantum yield at 355 nm, with polyAspartate-NaCl and polyArginine-HCl showing Φ_{355} around 0.0398 and 0.030, respectively. Among the amino acids, Aspartate-KCl, and Arginine showed similar Φ_{355} ~0.026, followed by Histidine (~0.017) and Arginine-HCl (~0.005). Interestingly, Symfoil PV2 showed the least Φ_{355} value of 0.0027. Since luminescence is related to quantum yield, we investigated their inter-dependence on luminescence sensitivity. We investigated the relationship between the luminescence output with the molar extinction coefficient (ε_{355}) and luminescence quantum yield (Φ_{355}). The slopes for luminescence intensity versus concentration were obtained from **Figure 3.8.** It was observed that the slope of luminescence was directly proportional to the product of the molar extinction coefficient and quantum yield, as well to the molar extinction coefficient alone, as shown in **Figure 3.13B** and **Figure 3.13C**, respectively. However, no correlation between ε_{355} and Φ_{355} was observed, which was quite striking, as shown in **Figure 3.13D**. Moreover, Symfoil PV2 was found to have the lowest Φ_{355}, although having the highest ε_{355} among all the samples. However, amino acids and poly-amino acids which had relatively lower ε_{355} as compared to Symfoil PV2, showed higher Φ_{355}. It thus signifies that the greater luminescence slopes observed for protein and peptides were observed mainly because of their higher values of ε_{355} rather than their lower quantum yield Φ_{355}, as compared to amino acids. Similarly, the low luminescence slopes observed in amino acids were primarily dictated by the lower molar absorptivity

Here, the luminescence quantum yields (Φ_{355} ~0.003-0.04) observed among all the samples were relatively low. The low efficiency of radiative recombination compared to the competing non-radiative pathways could justify the lower quantum yields. Similar reports of low quantum yield have also been reported among identical chromophores[45]. It is interesting to point out the fact that Symfoil PV2 being a protein, showed a relatively lower quantum yield than polypeptide or even amino acids. The only shortcoming in ProCharTS luminescence is related to their low quantum yield since around ~97% of the relaxation occurs by the non-radiative pathway. Thus, effective modulation of ProCharTS by harnessing these electrons to recombine radiatively would help develop ProCharTS as a better tool for studying different protein dynamics. Recently, various strategies to increase the quantum yield of non-conventional chromophores have been reported[168]. Such modulation of successfully improving the quantum

yield at longer wavelength spectral regions, i.e., above 400 nm, would be a monumental feat in the field of protein chemistry and *in vivo* studies.

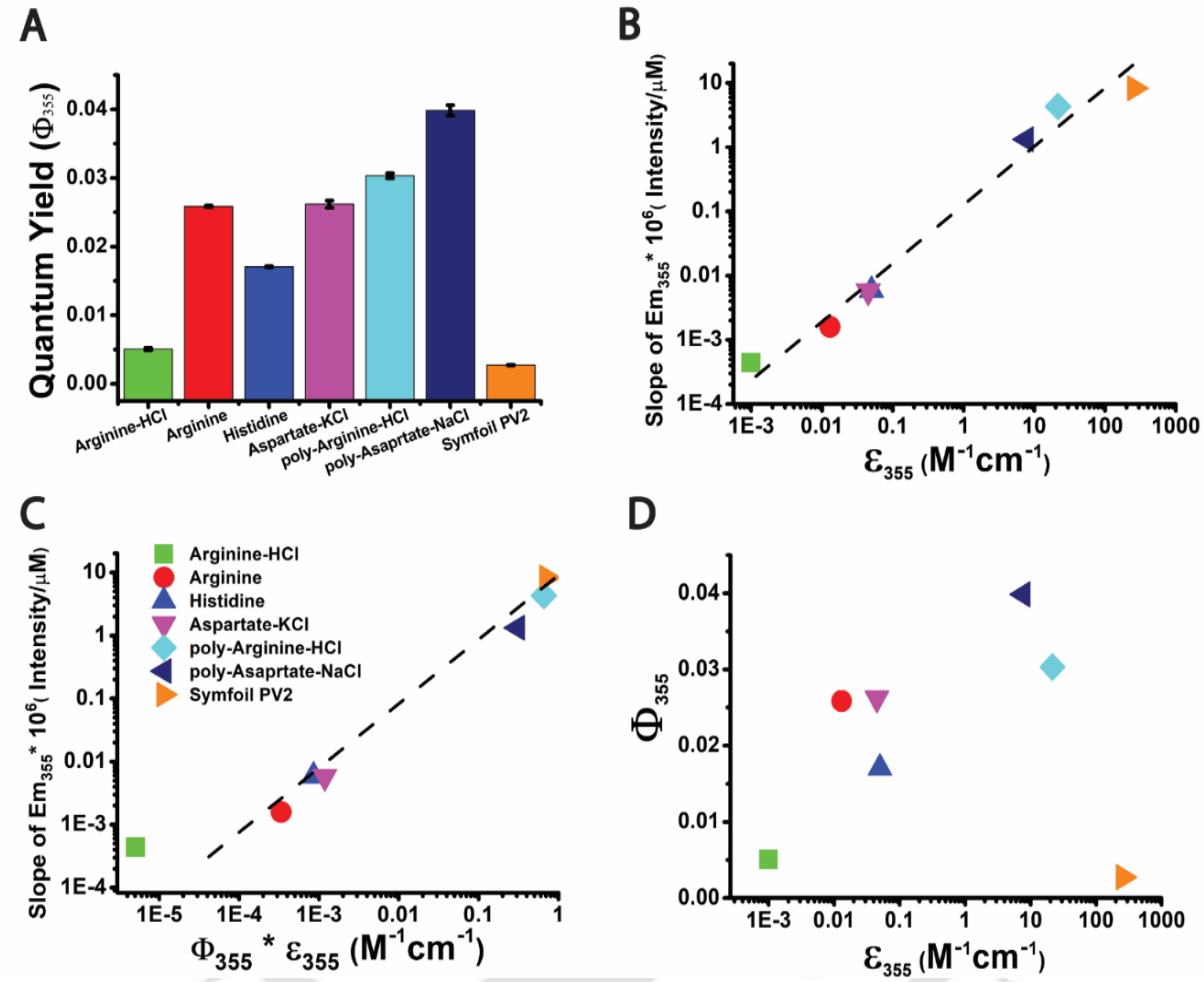


Figure 3.13. Luminescence Quantum yields of all samples excited at 355 nm (A); Correlation between the slope of luminescence integrated intensity/concentration plot with extinction coefficient at 355 nm (B); Correlation of emission slope with the product of Φ and extinction coefficient at 355 nm (C); and Correlation of Φ with extinction coefficient at 355 nm (D), for Arginine-HCl, Arginine, Histidine and Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 are shown.

3.2.5. Time-resolved luminescence intensity decay analysis

In the earlier sections, the interplay between ProCharTS absorbance and the steady state ProCharTS luminescence and their intrinsic properties like Stokes Shift, Quantum Yield, and emission maxima have been analyzed. Here, we tried to characterize the time-resolved luminescence decay parameters at 295 and 340 nm. The choice of the wavelength was based on the fact that conventional intrinsic chromophore like Trp is excited chiefly at 295 nm. Thus,

excitation at a similar wavelength could give us an idea about the relative timescale of ProCharTS decay kinetics compared to Trp. Moreover, 340 nm was used because most samples showed relatively higher luminescence emission intensity at 340 nm excitation. Furthermore, we extracted parameters from the luminescence intensity decay data by Non-linear least square (NLLS) and Maximum Entropy Method (MEM) based luminescence lifetime analysis as discussed below.

3.2.5.1. Non-linear least square (NLLS) analysis

The ProCharTS luminescence decay when excited at 295 and 340 nm, respectively, for all the samples were obtained and subsequently fitted. The fitted decays and their residuals are shown in **Figure 3.14** and **Figure 3.15**. Parameters extracted from the decay analysis of luminescence intensity at 295 nm excitation are shown in **Table 3.2** and **Figure 3.16 (A, C, E, G)**. A 3-exponential model fitted best for all the samples with three discrete lifetimes with varying amplitudes. Both 1 and 2-exponential fits were performed, but the fit did not converge properly. As evident from **Figure 3.14**, the residuals shown beneath every fitted decay appear random, signifying that decay data have been fitted properly in the 3-exponential model with χ_R^2 closer to 1.0 for all the samples.

All the amino acids, peptides, and Symfoil PV2 were characterized by a short luminescence lifetime component of τ_1 ~ 0.5-1.1 ns, followed by an intermediate lifetime of τ_2 ~ 2-4 ns and a longer lifetime of τ_3 ~ 6-11 ns. For poly-Aspartate-NaCl, however, all the lifetime components showed longer lifetime values than the rest. The contribution to the luminescence decay from the shorter component was most with α_1~ 0.7-0.8, followed by the intermediate lifetime component with α_2 ~ 0.1-0.3, and least for the longest lifetime component with $\alpha_3 < 0.1$. The mean luminescence lifetime τ_{mean} was least for Arginine-HCl and Symfoil PV2 (~0.9 ns), moderate for Arginine and poly-Arginine-HCl (~1.0-1.5 ns), followed by Histidine and poly-Aspartate-NaCl (~1.7 ns) and most for Aspartate-KCl (~2.1 ns).

The parameters extracted from the decay analysis of luminescence intensity at 340 nm excitation are shown in **Table 3.3** and **Figure 3.16** (**B, D, F, H**). Similarly, in this case, the 3-exponential model fitted best for all the samples with three discrete lifetimes with varying amplitudes and with better residuals with the χ_R^2 closer to 1.0 for all the samples. All the samples showed a shorter lifetime of ~ 1 ns with the most amplitude α_1 ~ 0.7, followed by an intermediate lifetime of ~ 2-3 ns with α_2 ~ 0.17-0.3 and a relatively long lifetime component

of ~ 6-13 ns with α_3 ~ 0.02-0.07. Interestingly, Aspartate-KCl showed a significantly lower α_1 value of 0.54 and a higher α_3 value of 0.17 than the rest of the samples. The τ_{mean} was least for

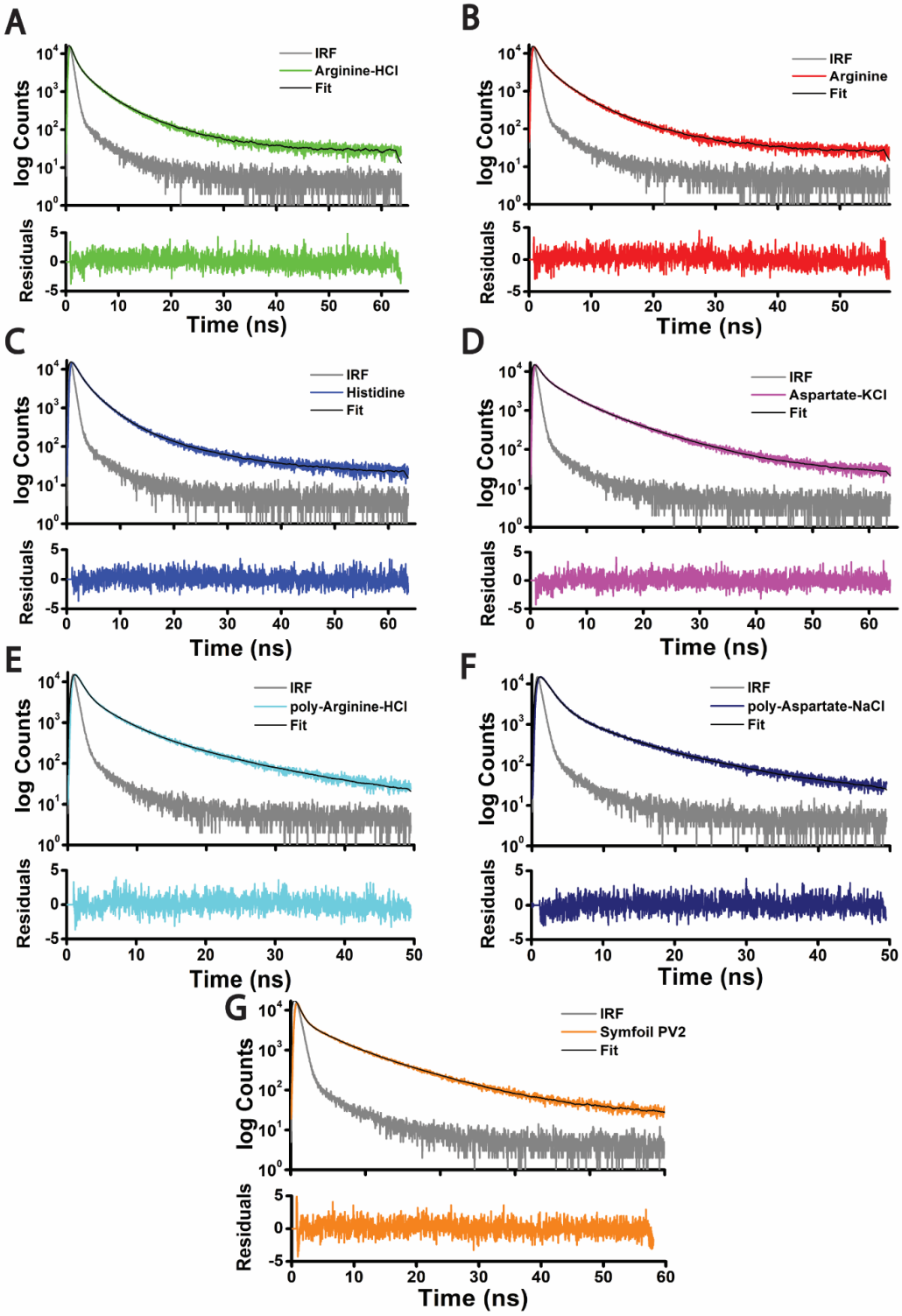


Figure 3.14. Luminescence Intensity Decay of Arginine-HCl (A), Arginine (B), Histidine (C) and Aspartate-KCl (D), poly-Arginine-HCl (E), poly-Aspartate-NaCl (F), Symfoil PV2 (G) when excited at 295 nm is shown here along with their respective residuals.

	τ_1(ns)	τ_2(ns)	τ_3(ns)	α_1	α_2	α_3	τ_{mean}(ns)	χ_R^2
Arginine-HCl	0.48(0.027)	2.16(0.065)	6.26(0.050)	0.807	0.162	0.031	0.92(0.043)	1.26
Arginine	0.54(0.004)	2.28(0.022)	6.20(0.091)	0.752	0.213	0.035	1.10(0.003)	1.14
Histidine	1.00(0.010)	2.94(0.015)	10.23(0.309)	0.691	0.291	0.018	1.73(0.002)	1.07
Aspartate-KCl	0.68(0.013)	3.49(0.038)	9.22(0.054)	0.732	0.246	0.083	2.10(0.025)	1.01
Poly-Arginine-HCl	0.76(0.016)	3.23(0.055)	10.03(0.048)	0.796	0.171	0.033	1.48(0.008)	1.13
Poly-Aspartate-NaCl	1.15(0.012)	4.18(0.114)	11.43(0.108)	0.869	0.104	0.027	1.76(0.003)	1.10
Symfoil PV2	0.40(0.007)	3.03(0.113)	7.69(0.219)	0.869	0.086	0.045	0.90(0.049)	1.08

Table 3.2. Parameters obtained from 3 exponential-fit for all the samples at excitation wavelength 295 nm. The individual lifetime components τ_i and mean lifetime τ_{mean} values obtained from 3 exponential-fit are averaged for at least 2 independent measurements and shown here, while the values in the parenthesis represent their standard deviation. The amplitude values α_i and the reduced Chi Square χ^2 represent the best fit data from the 2 or more measurements for each sample.

	τ_1(ns)	τ_2(ns)	τ_3(ns)	α_1	α_2	α_3	τ_{mean}(ns)	χ_R^2
Arginine-HCl	1.05(0.009)	3.77(0.048)	10.55(0.497)	0.730	0.253	0.017	2.00(0.030)	1.08
Arginine	1.07(0.065)	3.34(0.224)	9.36(1.084)	0.664	0.298	0.038	1.96(0.003)	1.02
Histidine	0.76(0.053)	2.38(0.146)	6.18(0.228)	0.665	0.281	0.054	1.54(0.038)	1.09
Aspartate-KCl	0.65(0.058)	2.83(0.169)	7.85(0.132)	0.546	0.284	0.170	2.52(0.068)	1.07
Poly-Arginine-HCl	0.76(0.005)	2.78(0.125)	7.78(0.228)	0.751	0.176	0.073	1.60(0.011)	1.13
Poly-Aspartate-NaCl	1.11(0.018)	3.18(0.210)	8.56(0.477)	0.754	0.182	0.063	2.01(0.008)	1.06
Symfoil PV2	0.69(0.039)	3.18(0.190)	13.03(0.717)	0.763	0.186	0.051	1.82(0.048)	1.01

Table 3.3. Parameters obtained from 3 exponential-fit for all the samples at an excitation wavelength of 340 nm. The individual lifetime components τ_i and mean lifetime τ_{mean} values obtained from 3 exponential-fit are averaged for at least 2 independent measurements and shown here, while the values in the parenthesis represent their standard deviation. The amplitude values α_i along with the reduced Chi Square χ^2 represent the best fit data from the 2 or more measurements for each sample.

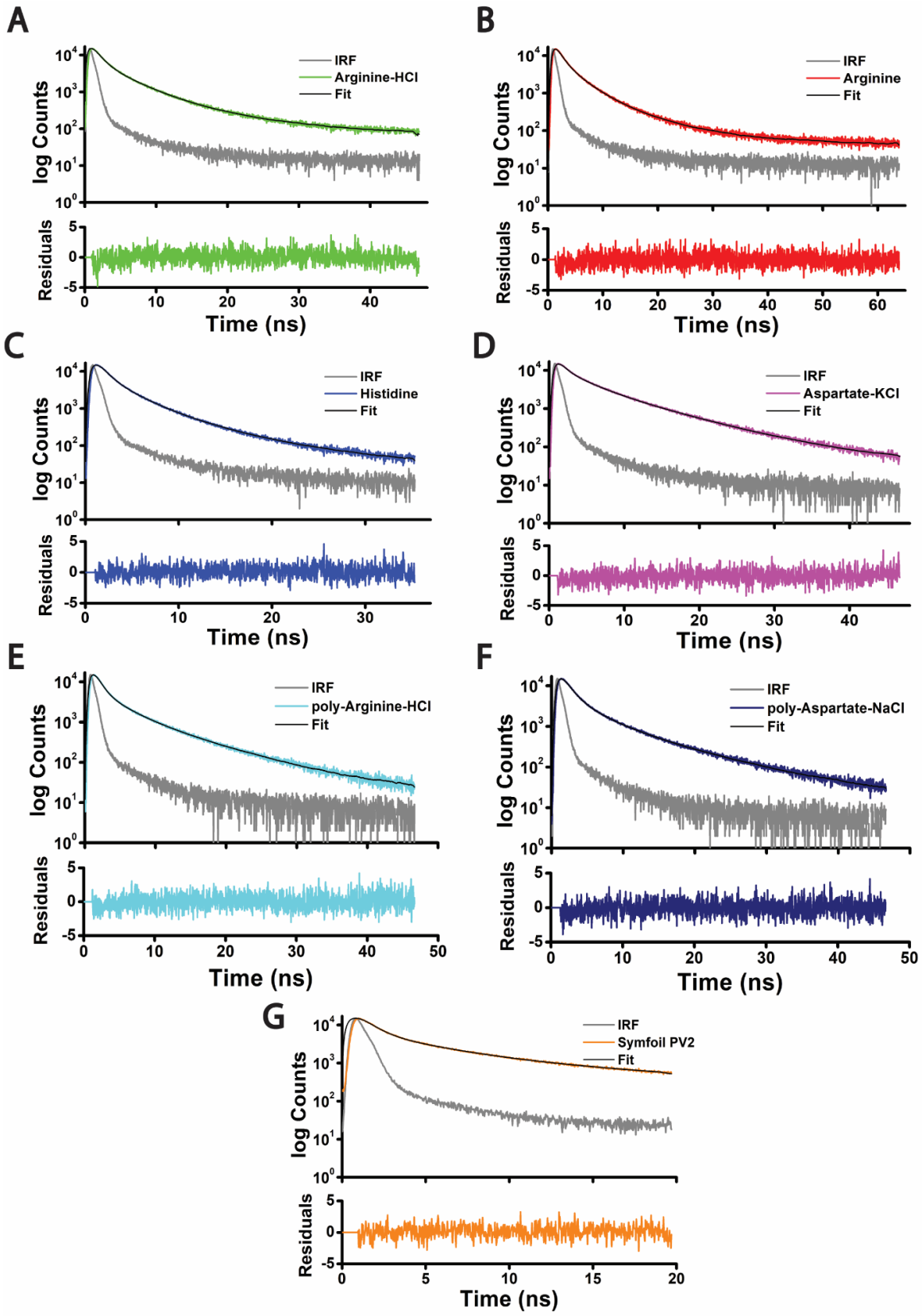


Figure 3.15. Luminescence Intensity Decay of Arginine-HCl (A), Arginine (B), Histidine (C) and Aspartate-KCl (D), poly-Arginine-HCl (E), poly-Aspartate-NaCl (F), Symfoil PV2 (G) when excited at 340 nm is shown here along with their respective residuals.

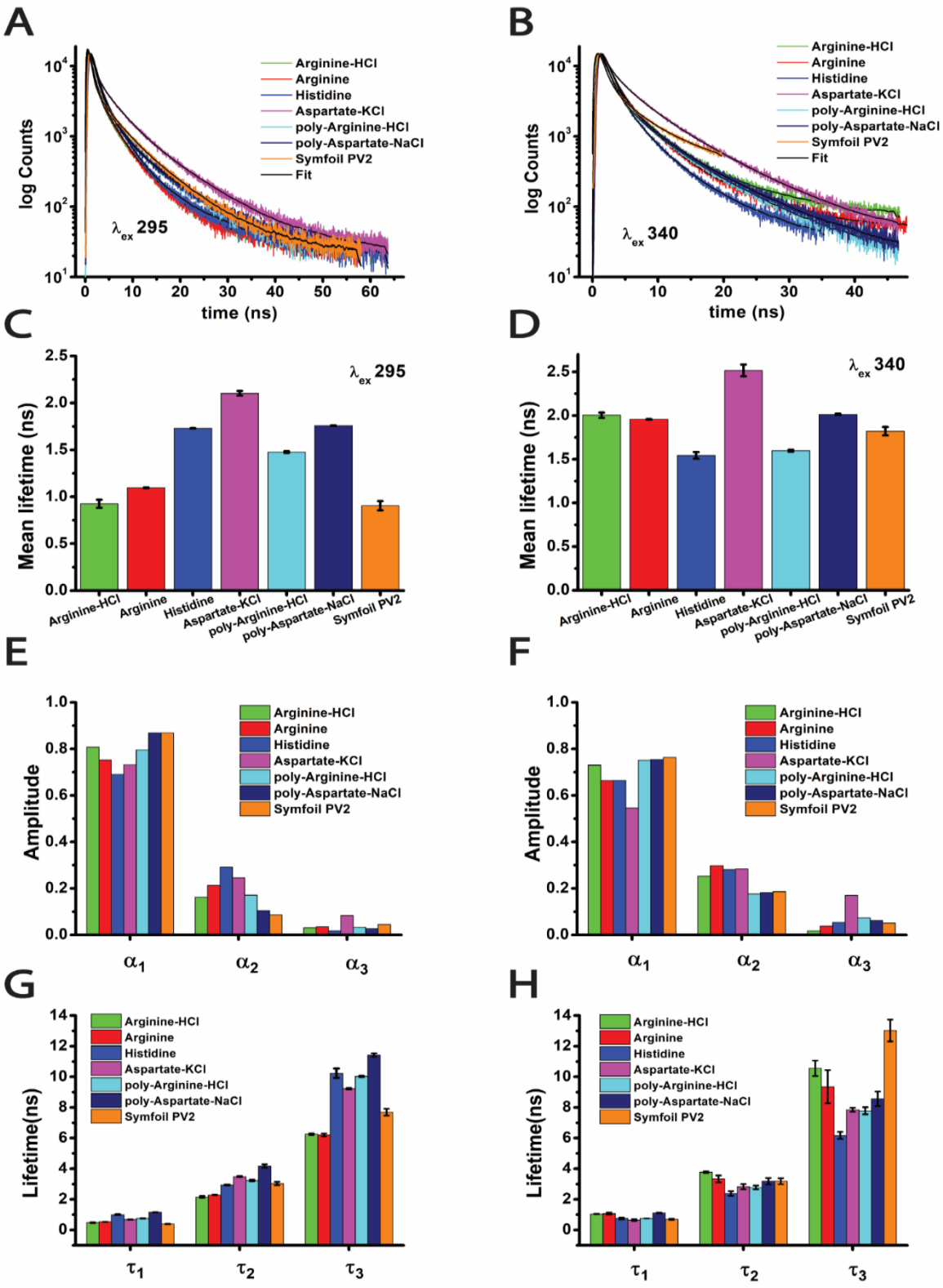


Figure 3.16. Luminescence Intensity Decay and the 3-exponential fitted data using Peri software at excitation 295 nm and 340 nm for all the samples: Arginine-HCl, Arginine, Histidine and Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 were shown in panel A and B. Subfigures C, E, and G show the mean luminescence lifetime, the amplitude of individual lifetimes and individual lifetime components when excited at

295 nm respectively, and subfigures D, F, H indicate the same for 340 nm respectively. All the lifetime values shown here are averaged from 2 independent experiments, and their respective amplitude values represent their best fit.

Histidine and poly-Arginine-HCl with values of ~ 1.5-1.6 ns, followed by Symfoil PV2, Arginine, Arginine-HCl, poly-Aspartate-NaCl showing moderate values of ~ 1.8-2.0 ns, and highest for Aspartate-KCl at ~ 2.5 ns. Also, τ_{mean} at 340 nm showed higher or similar values than at 295 nm for most of the samples, with the only exception being Histidine. Such ultrafast decay kinetics has been observed in Lysine rich proteins and amino acids[30,169,170].

3.2.5.2. Maximum Entropy Method (MEM) analysis

Here, the MEM was used to alternately extract the luminescence lifetime distribution from the fitted intensity decays at both the 295 and 340 nm excitation. The MEM fit residuals for the 295 and 340 nm excitation for all the samples are shown in **Figure 3.17** and **Figure 3.18**, respectively. All the samples displayed a trimodal luminescence lifetime distribution with varying amplitudes as shown in **Figure 3.19**. A common trend of longer lifetimes with lower amplitudes while shorter lifetimes had higher amplitudes was quite evident in most samples. It could be observed that the luminescence lifetime distribution for the short component τ_1 at 295 nm was more shifted towards shorter lifetime values as compared to 340 nm. However, the lifetime distributions for the two other longer components (τ_2 and τ_3) were consistent. Similar behaviour could also be seen in **Figure 3.16** from the discrete fit analysis. The luminescence distribution at 295 nm was found to be more conserved for each lifetime than at 340 nm. The spread for the distribution for an individual lifetime at 295 nm (τ_1 < 0.5 ns, τ_2 ~ 1-3 ns, τ_3 ~ 3-8 ns) was found to be less as compared to that of 340 nm (τ_1 < 1 ns, τ_2 ~ 1-6 ns, τ_3 ~ 3-14 ns).

The origin of these three different luminescence lifetimes is proposed to arise from the three separate charge recombination pathways of electrons coming down from a single excited state to the ground state. The shorter luminescence lifetime (<1 ns) is thought to be arising from the direct recombination of electrons from the less to the more electronegative species. However, the longer lifetime component is proposed to originate from delayed recombination. In delayed recombination, the excited state electron from the less electronegative species presumably hops first onto the protein surface and subsequently to more electronegative species in the protein matrix. Although the MEM distribution supplements the Discrete fit analysis, the broad distribution of the shorter lifetime component observed in MEM sheds light

on the possibility of multiple additional short lifetime subcomponents that were not resolved in the Discrete analysis.

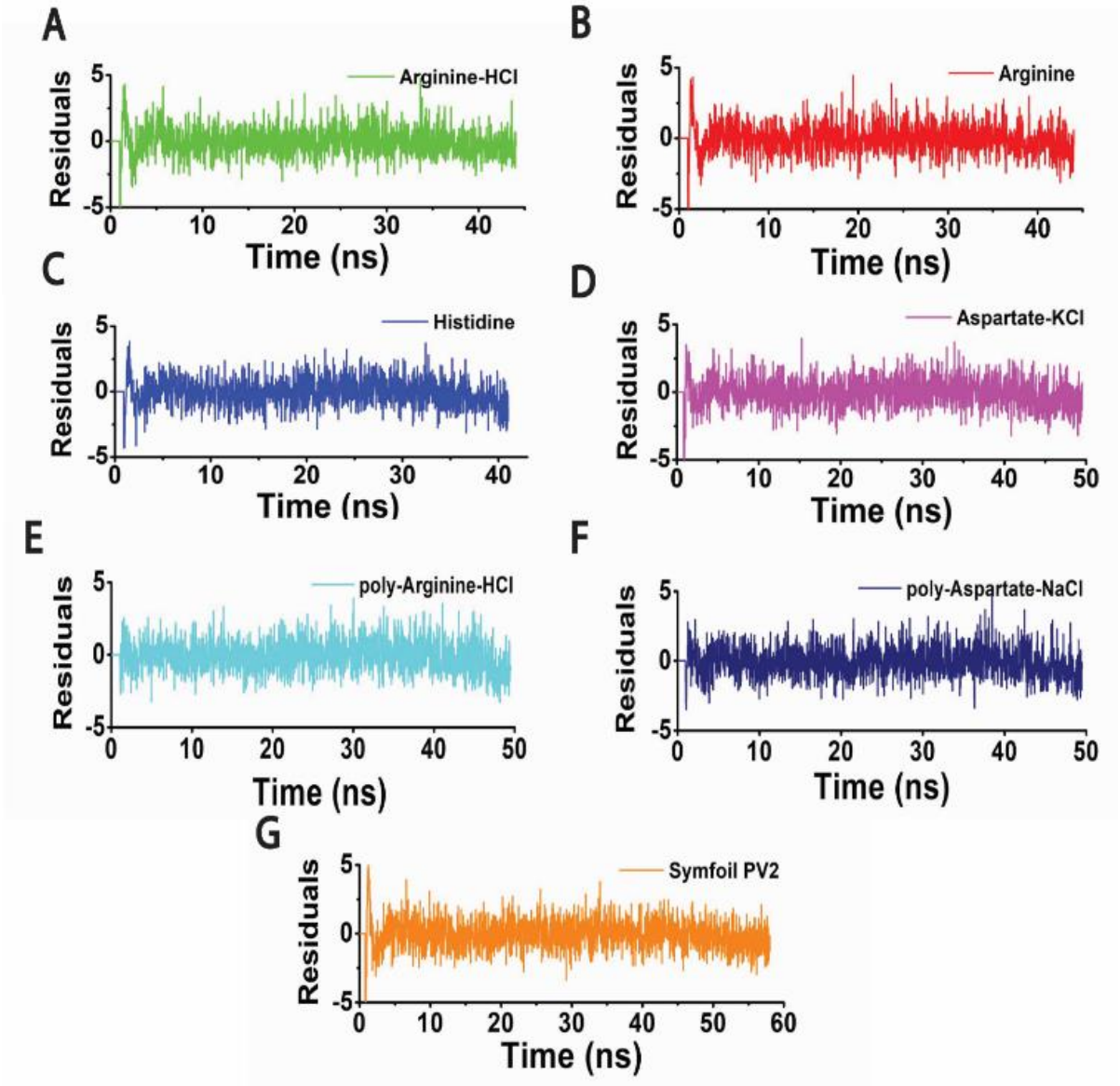


Figure 3.17. Residuals of luminescence lifetime decay by MEM analysis at 295 nm excitation are shown here. The panel represents residuals for fitted data of Arginine-HCl (A), Arginine (B), Histidine (C), Aspartate-KCl (D), poly-Arginine-HCl (E), poly-Aspartate-NaCl (F), and Symfoil PV2 (G).

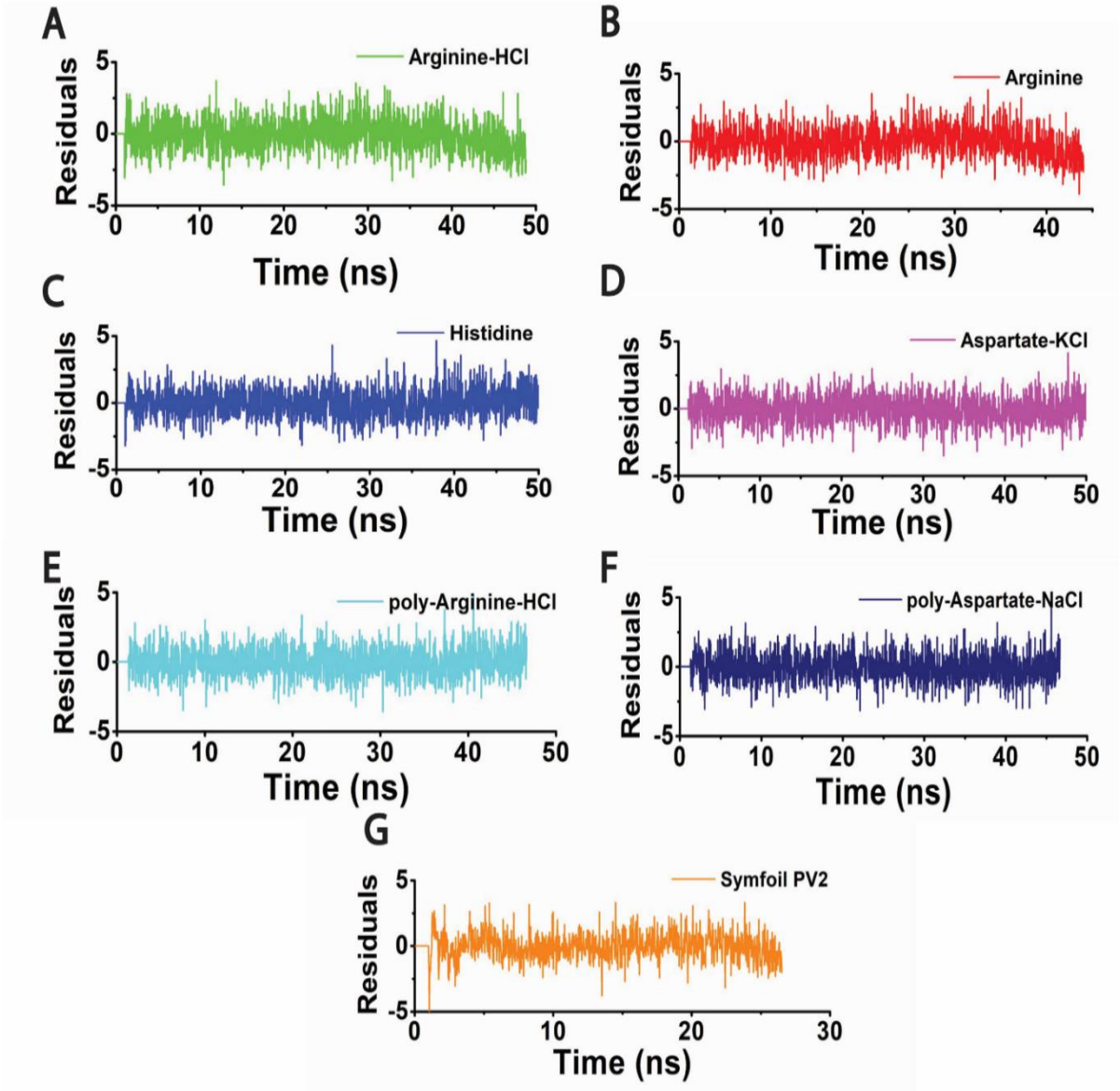


Figure 3.18. Residuals of luminescence lifetime decay by MEM analysis at 340 nm excitation are shown here. The panel represents residuals for fitted data of Arginine-HCl (A), Arginine (B), Histidine (C), Aspartate KCl (D), poly-Arginine-HCl (E), poly-Aspartate-NaCl (F), and Symfoil PV2 (G).

Similarly, studies on the origin of multiple luminescence lifetime components have been reported in cases of amyloid oligomers and Lysine-rich proteins[30,32]. However, it is for the first time that such common characteristics have been found in Arginine/Aspartate rich peptide/protein systems. Thus, ProCharTS has the potential to characterize various protein systems along with their interaction with other biomolecules. This may open the door to newer applications and possibilities in this field of non-conventional chromophores. Recently, different cell-penetrating Arginine polypeptides have successfully delivered many biologically active macromolecules[171]. Although, more work needs to be done to improve the luminescence

quantum yield by channelling the charge-separated excited states to recombine radiatively so as to enhance the applicability of ProCharTS.

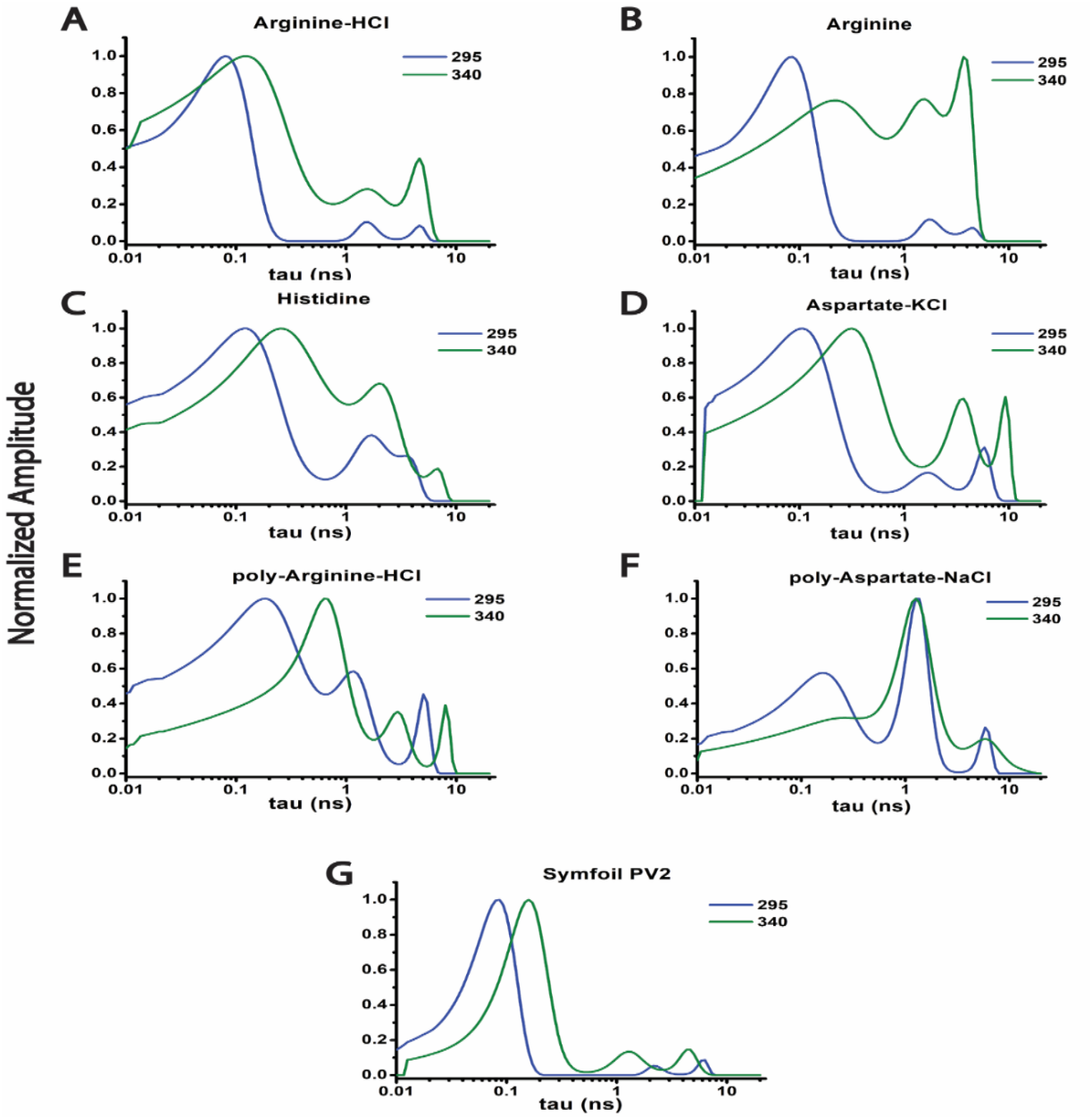


Figure 3.19. MEM luminescence lifetime distribution on excitation at 295 nm and 340 nm for Arginine-HCl, Arginine, Histidine, Aspartate-KCl, poly-Arginine-HCl, poly-Aspartate-NaCl, and Symfoil PV2 are shown in subfigures A to G, respectively

3.3. Conclusions

- The ProCharTS absorbance and luminescence features are highly conserved among charged amino acids, peptides, and proteins lacking Lysine. ProCharTS is characterized by a long absorbance tail extending till 800 nm, significant Stokes shift, low quantum yield, the origin of three luminescence lifetime components, and emission maximum increasing with increased excitation wavelength.

- The origin of ProCharTS is evident as long as charge transfer occurs among proximal charged residues, irrespective of the type of charged amino acids present in such systems.

- The emergence of multiple excitation peaks and varied luminescence lifetime indicate that a heterogeneous population of chromophores manifests the ProCharTS phenomenon.

- The significantly higher contribution from the lower luminescence lifetime values suggests that the charge recombination phenomenon is quite fast. The lower quantum yield further indicates that most of the photoexcitation relaxes via non-radiative transfer.

- The strong correlation between the slope of integrated luminescence intensity with the molar extinction coefficient and also with the product of extinction coefficient and quantum yield points out that it is the extinction coefficient that dictates the ProCharTS luminescence, rather than any other factor.

Chapter 4

Role of charged amino acids in sullying the fluorescence of Tryptophan in monomeric protein

Chapter 4

4.1. Introduction

Tryptophan (Trp) is a naturally occurring amino acid popularly used as an intrinsic fluorescent probe to monitor protein conformation and dynamic changes. The indole group in Trp is mainly responsible for its emissive properties. The extreme sensitivity of the Trp fluorescence emission towards environmental changes near the indole group makes Trp a novel intrinsic probe to study protein interaction with other biomolecules. The change in the emission is reflected in its emission maxima (λ_{max}), fluorescence quantum yields (Φ) and its fluorescence lifetimes components. Proteins or peptides with a single Trp residue typically display fluorescence decays that are fit by sums of two to four exponential terms, with varying amplitudes and lifetime components. Such a single Trp system has been extensively used due to the ease of analysis of the multi-exponential changes upon conformational changes. However, the origin of such multi-exponentiality has been controversial. In fact, NATA, a tryptophan analogue with the same indole group, displays one fluorescence lifetime component in water. Various opinions about the role of electron transfer towards the origin of the fluorescence in the indole group inside the protein matrix have been hypothesised[52,63,172–175]. Recently, ProCharTS is thought to play a role in sullying the Trp fluorescence[67]. Lysine was found to have the most significant contribution to contaminating the indole fluorescence, presumably because of its higher absorptive and luminescence properties among the charged amino acids. The charge recombination process observed in ProCharTS through direct or delayed pathways could further explain the origin of Tryptophan's multiple fluorescence lifetimes.

In this chapter, I mainly focussed on the role of charged amino acids other than Lysine and Glutamate in contaminating the Trp fluorescence. A protein system with a single Tryptophan, surrounded by Arginine, Aspartate and Histidine residues, was mimicked by the tryptophan analogue, NATA, in the presence of these charged amino acids. Subsequently, the changes in its luminescence lifetime were analysed using the nonlinear least square (NLLS) and Maximum Entropy Method (MEM). Moreover, a similar study has been carried out in the presence of Symfoil PV2, which is rich in Arginine, Aspartate and Histidine. In Symfoil PV2, an even more realistic scenario is reproduced due to the close 3D proximity of NATA to the terminal head groups of the charged amino acids and to the backbone of the peptide bond inside the protein matrix.

4.2. Results and Discussions

4.2.1. Quenching of NATA fluorescence by charged amino acids

The steady-state fluorescence of NATA is measured in the presence of increasing amino acid concentration. NATA shows an extinction coefficient of 5500 $M^{-1}cm^{-1}$ at 280 nm[151] and a quantum yield of around 0.14 at 295 nm[2]. Since both the quantum yield (0.017-0.026) and the extinction coefficients (0.12-0.27) of Arginine, Aspartate and Histidine residues at 295 nm are around 10 and 1000 times lower than that of NATA, respectively, a 10000-fold excess of amino acids was used to have a noticeable impact on its fluorescence. Also, such a high concentration would mimic a protein-like environment where charged amino acids are close to the indole ring in NATA. For all the cases, the fluorescence of NATA was quenched with increasing concentration, as shown in **Figure 4.1**.

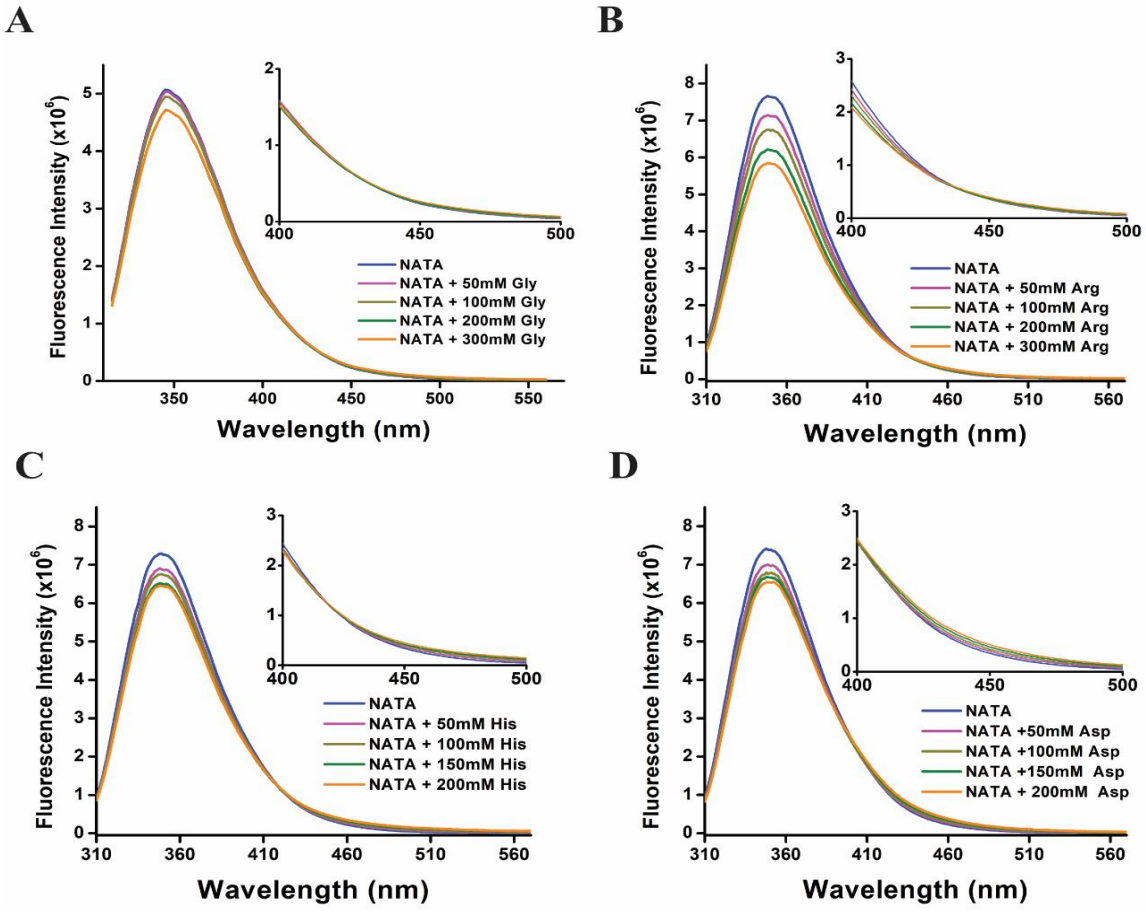


Figure 4.1: Steady-state fluorescence emission of NATA in the presence of charged amino acids at 295 nm excitation. Subfigures A, B, C, and D show the fluorescence of NATA in the presence of Glycine, Arginine, Histidine and Aspartate, respectively. Insets show the corresponding fluorescence emission from 400-500 nm

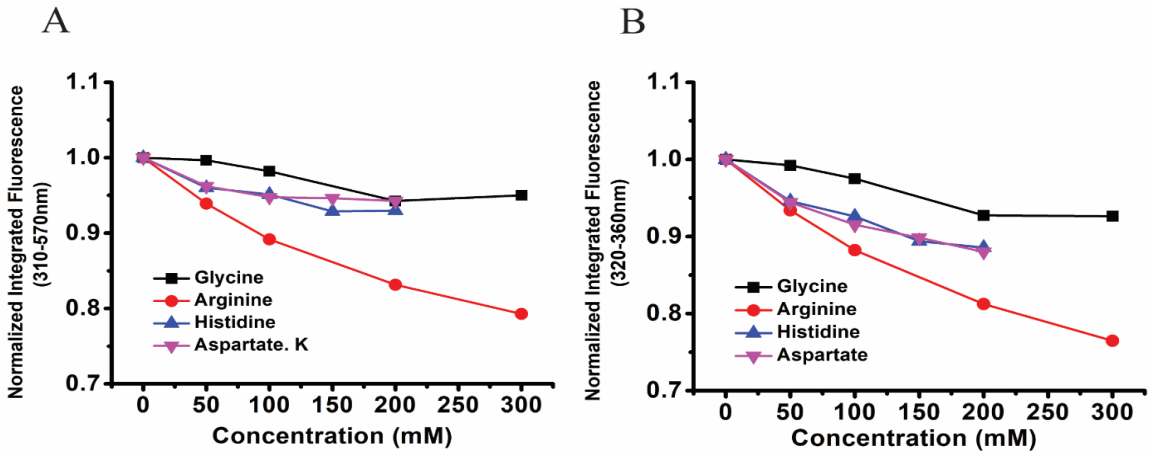


Figure 4.2: Area (normalised) under NATA emission spectrum in the presence of different amino acids for A) emission between 310-570 nm and B) emission between 320-360 nm.

We wanted to quantify the changes, and thus we considered two spectra regions: one from 310-570 nm and the other from 320-360 nm. It has been observed that for both these regions, the decrease in integrated fluorescence of NATA was consistent, and it was found to be most for Arginine, followed by Aspartate, Histidine, and the least for Glycine. Here, Glycine, which does not have a charged side chain, showed the least quenching of NATA, which was quite expected. Thus, the amino acid's charged side chain plays a significant role in quenching NATA fluorescence. Also, one important observation, as seen in **Figure 4.1**, was that as the fluorescence was quenched, the spectra shifted more towards the red region, as decreased fluorescence at 320-360 nm resulted in an increase of fluorescence in the 420-570 nm region. Similarly, the decrease in fluorescence integrated area in **Figure 4.2A** was less than that in **Figure 4.2B**. Thus, in the presence of charged amino acids, a common trend of noticeable concentration-dependent decrease of fluorescence intensity in the spectral region (320-360 nm) of NATA emission was observed with a modest increase in emission on the red side of the spectrum contaminating the indole fluorescence. This rise in luminescence beyond 430 nm was thought to arise from individual amino acids emission as the charged amino acids have their λ_{max} around 400-420 nm when excited at 295 nm, as shown in *Chapter 3*.

4.2.2. Fluorescence intensity decay of NATA in the presence of charged amino acids

4.2.2.1. Non-linear Least Square (NLLS) analysis

Time-resolved fluorescence intensity of NATA in the presence of different charged amino acids was recorded at λ_{ex} 295 nm. The decay was fitted to a one exponential decay model as

NATA, when excited at 295 nm, was expected to follow the single exponential decay. Although the fluorescence intensity decays of NATA could be fitted in one-exponential with lower concentrations, but not with increasing concentration of charged amino acids, as can be seen from the residual distribution [**Figure 4.3, 4.4, 4.5 (A and C)**]. However, when the same intensity decays were fitted to the two-exponential model, the fits were acceptable with better residuals [**Figure 4.3, 4.4, 4.5 (B and D)**]. This pattern was consistent for all the charged amino acids.

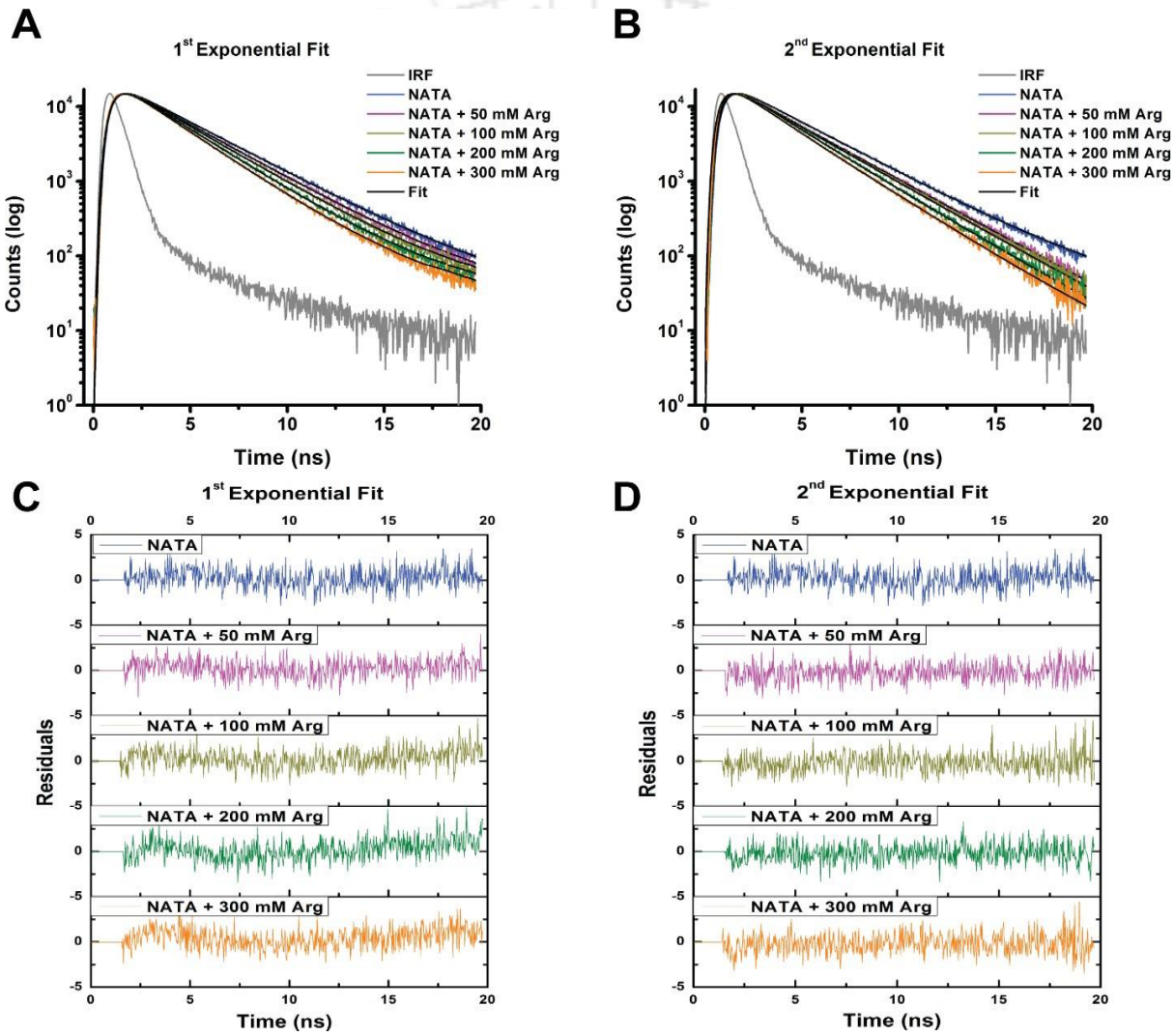


Figure 4.3: Fluorescent Intensity Decay of NATA at 295 nm excitation in the presence of different concentrations of Arginine. Subfigures A, C and B, D show the decay of NATA and the fitted residuals for the same data set fitted in 1 and 2 exponential models using the NLLS method, respectively.

	1 exponential-fit		2 exponential-fit					
	τ_{mean} (ns)	χ_R^2	α_1	α_2	τ_1 (ns)	τ_2 (ns)	τ_{mean} (ns)	χ_R^2
NATA	3.05 (0.006)	1.143	1	---	3.05 (0.006)	---	3.05 (0.006)	1.143
NATA+50 mM Arg	2.84 (0.026)	1.005	0.706	0.294	3.17 (0.013)	1.99 (0.051)	2.85 (0.014)	1.004
NATA+100 mM Arg	2.70 (0.016)	1.109	0.647	0.353	3.09 (0.037)	2.04 (0.004)	2.72 (0.019)	1.085
NATA+200 mM Arg	2.51 (0.005)	1.297	0.317	0.683	3.17 (0.058)	2.17 (0.008)	2.52 (0.019)	1.047
NATA+300 mM Arg	2.34 (0.005)	1.302	0.104	0.896	3.62 (0.054)	2.21 (0.014)	2.37 (0.010)	1.079

Table 4.1: Parameters obtained from 1 and 2 exponential NLLS fit for NATA in the presence of Arginine

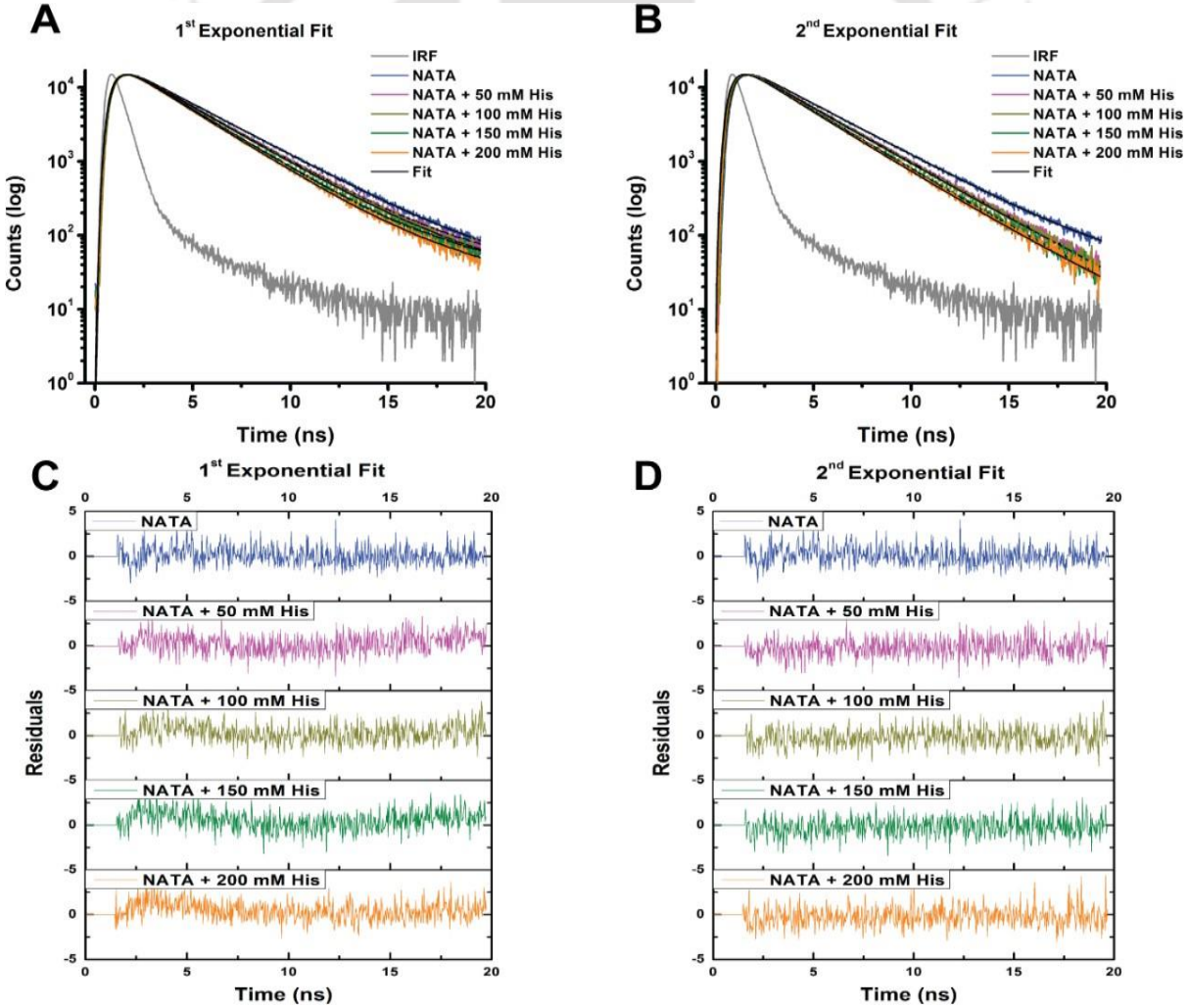


Figure 4.4: Fluorescent Intensity Decay of NATA at 295 nm excitation in the presence of different concentrations of Histidine. Subfigures A, C and B, D show the decay of NATA and the fitted residuals for the same data set fitted in 1 and 2 exponential models using the NLLS method, respectively.

	1 exponential-fit		2 exponential-fit					
	τ_{mean} (ns)	χ_R^2	α_1	α_2	τ_1 (ns)	τ_2 (ns)	τ_{mean} (ns)	χ_R^2
NATA	2.95 (0.025)	1.016	1	---	2.95 (0.024)	---	2.95 (0.024)	1.016
NATA+50 mM His	2.75 (0.006)	1.232	0.759	0.241	3.06 (0.024)	1.99 (0.135)	2.77 (0.001)	1.144
NATA+100 mM His	2.68 (0.018)	1.118	0.495	0.505	3.19 (0.006)	2.20 (0.016)	2.68 (0.003)	1.072
NATA+150 mM His	2.56 (0.003)	1.448	0.324	0.676	3.25 (0.000)	2.26 (0.009)	2.58 (0.0007)	1.08
NATA+200 mM His	2.48 (0.005)	1.267	0.131	0.869	3.79 (0.113)	2.32 (0.012)	2.49 (0.012)	1.078

Table 4.2: Parameters obtained from 1 and 2 exponential NLLS fit for NATA in the presence of Histidine.

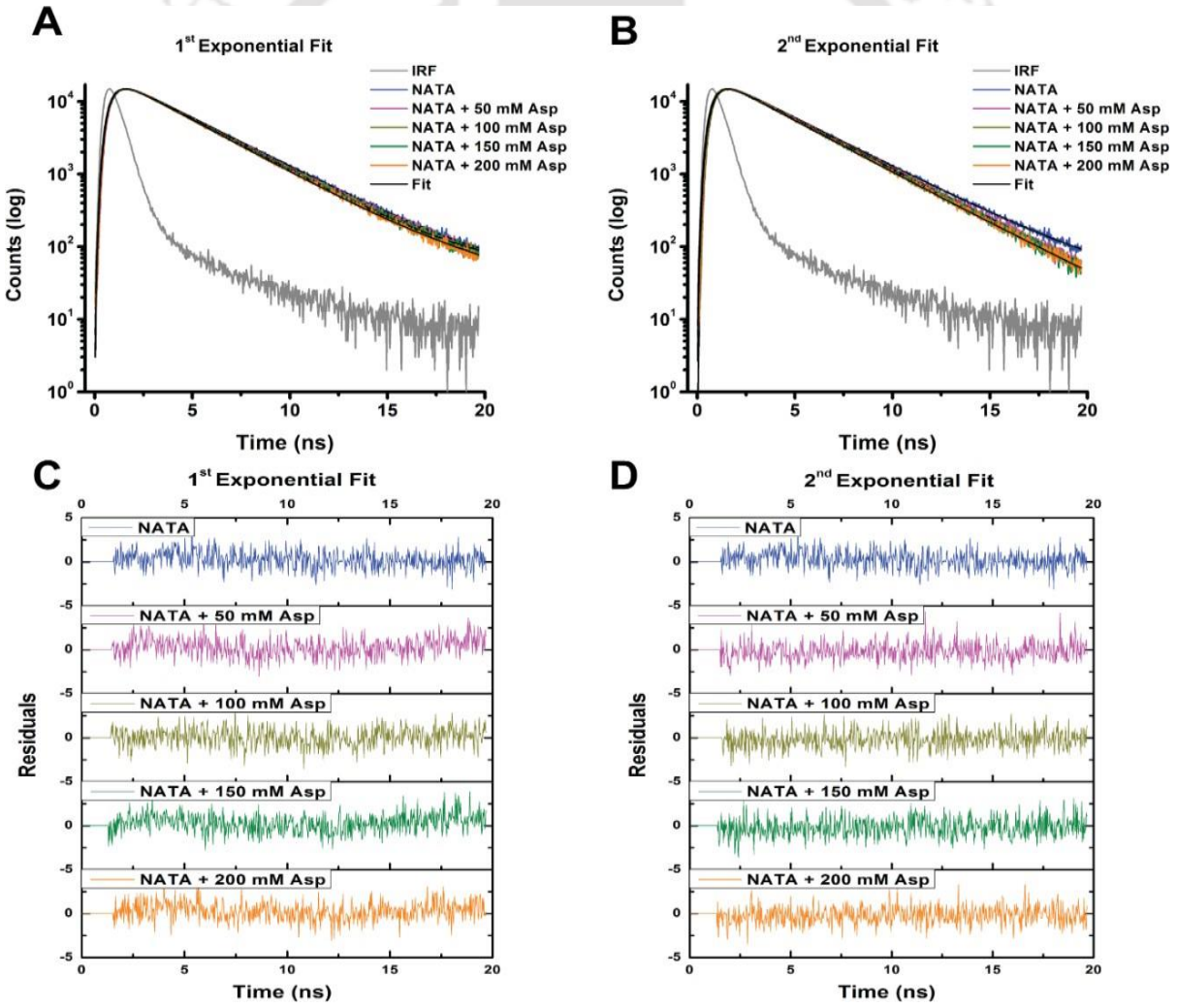


Figure 4.5: Fluorescent Intensity Decay of NATA at 295 nm excitation in the presence of different concentrations of Aspartate. Subfigures A, C and B, D show the decay of NATA and the fitted residuals for the same data set fitted in 1 and 2 exponential models using the NLLS method, respectively.

	1 exponential-fit		2 exponential-fit					
	τ_{mean} (ns)	χ_R^2	α_1	α_2	τ_1 (ns)	τ_2 (ns)	τ_{mean} (ns)	χ_R^2
NATA	3.04 (0.026)	1.034	1	---	3.04 (0.026)	---	3.04 (0.026)	1.034
NATA+50 mM Asp	2.94 (0.011)	1.167	0.698	0.302	3.31 (0.013)	2.13 (0.068)	2.96 (0.014)	1.037
NATA+100 mM Asp	2.93 (0.006)	1.061	0.67	0.33	3.28 (0.019)	2.21 (0.098)	2.93 (0.017)	1.048
NATA+150 mM Asp	2.86 (0.011)	1.139	0.601	0.399	3.30 (0.010)	2.24 (0.001)	2.89 (0.009)	1.038
NATA+200 mM Asp	2.84 (0.008)	1.091	0.462	0.538	3.37 (0.066)	2.34 (0.059)	2.85 (0.007)	1.011

Table 4.3: Parameters obtained from 1 and 2 exponential NLLS fit for NATA in the presence of the Aspartate

We, thereby plotted the χ_R^2 values from the fit for each amino acid, as shown in **Figure 4.6.** The χ_R^2 values deviated from 1 when fitted in one exponential, whereas it stayed close to 1 in the two-exponential model, thus justifying the multi-exponentiality seen in Trp fluorescence.

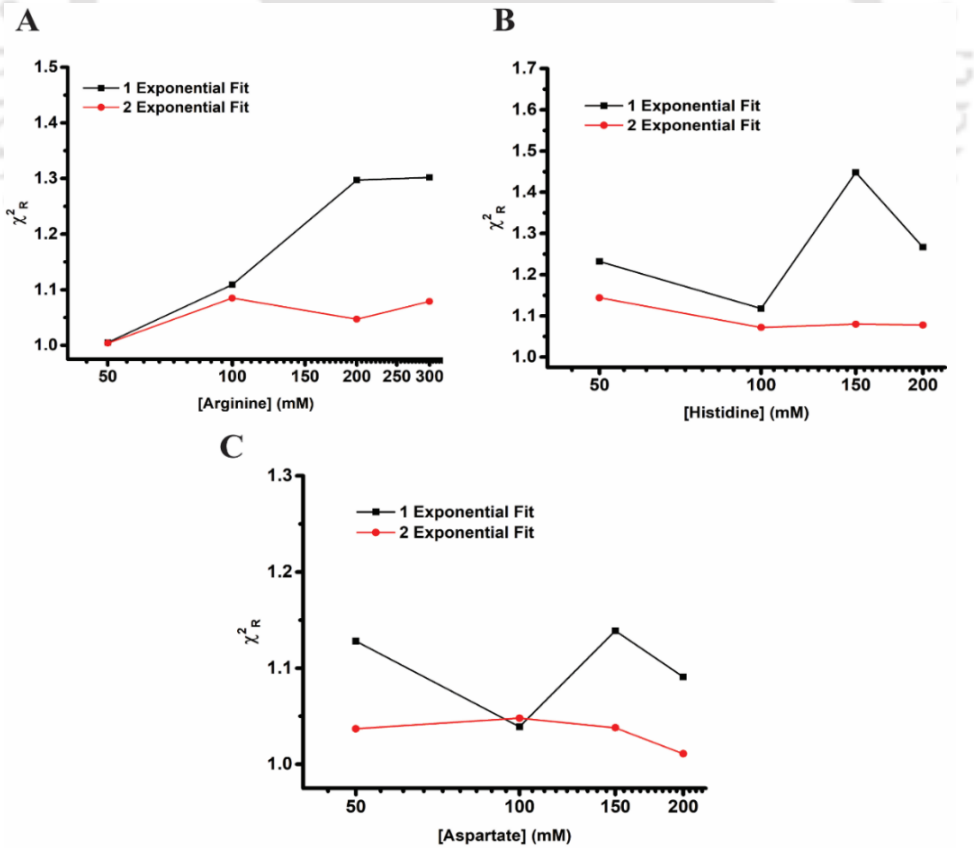


Figure 4.6: Reduced Chi-Square of the fluorescence intensity decay of NATA. Subfigures A, B, and C show the χ_R^2 values from both exponential fit for NATA in presence of Arginine, Histidine and Aspartate, respectively.

From the two-exponential fitted data, we obtained the mean lifetime for all the concentrations for each amino acid. We found that the mean lifetime 'τ' decreased with increasing concentration suggesting quenching of the NATA in the presence of amino acids. We thus calculated the τ_0/τ for all the amino acids with increasing concentration. Distinctive slopes were seen for each of these amino acids. A linear Stern-Volmer plot was generated, suggesting dynamic quenching of NATA. NATA was uniformly accessible to the quencher in each case, as seen from the linear Stern-Volmer plot (**Figure 4.7**), with Arginine having the highest slope followed closely by Histidine and then least being Aspartate. Moreover, the bimolecular quenching constant from the slopes was calculated to be around 0.3 x 10^9 $M^{-1}s^{-1}$ for Arginine, followed by Histidine and least for Aspartate (~0.1 x 10^9 $M^{-1}s^{-1}$).

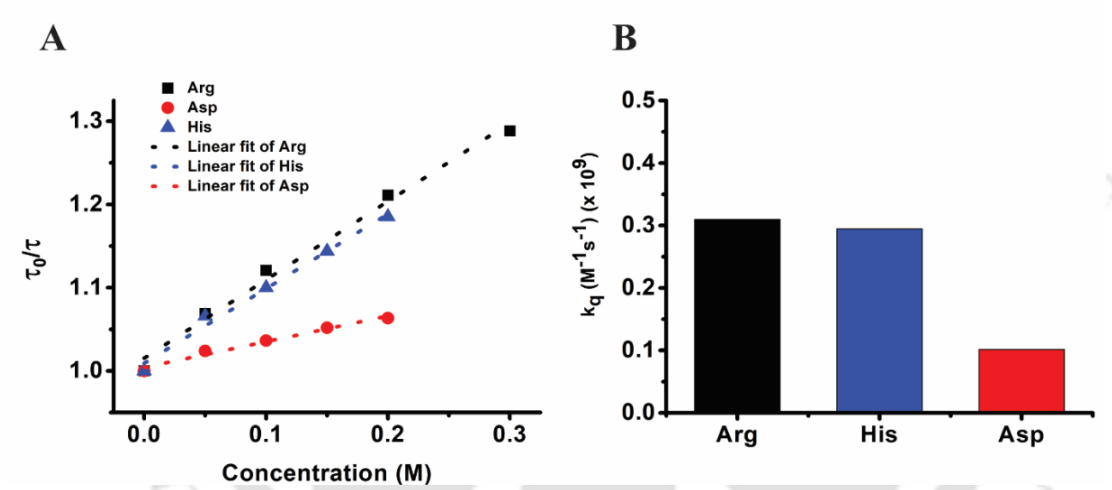


Figure 4.7: Subfigure A shows the Stern Volmer plot for the quenching of NATA in the presence of different amino acids. The dotted lines represent the individual fits. Subfigure B shows the bimolecular quenching constant for the same.

The fluorescence lifetime of NATA is generally found to be within 2.8-3.1 ns and is known to have only a one-lifetime component. However, with the increasing concentration of charged amino acids, the emergence of the shorter fluorescence lifetime component τ_1, as evident from the 2-exponential fit shown in **Tables 4.1, 4.2** and **4.3**. The decrease in the mean lifetime was most for Arginine, where it changed from 3.05 ns to 2.37 ns, followed by Histidine, which changed to 2.49 ns and least value of 2.85 ns for Aspartate, as can be observed in **Figure 4.8**. This decrease in the mean lifetime of NATA when excited at 295 nm resulted from the emergence of the shorter fluorescence lifetime component. The steady decline in the values of mean lifetime closely resembles the pattern observed in the steady-state quenching analysis.

Thus, the time-resolved analysis complemented the steady-state data, suggesting quenching of NATA in the presence of amino acids.

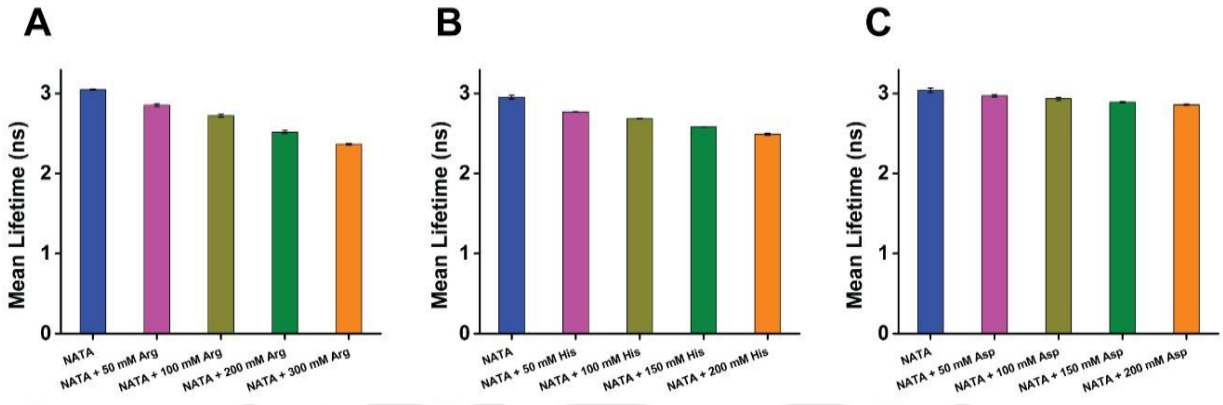


Figure 4.8: Mean fluorescence lifetime of NATA in the presence of increasing amino acid concentration. Panel A, B, and C show the mean fluorescence lifetime of NATA in the presence of Arginine, Histidine and Aspartate, respectively.

4.2.2.2. Maximum Entropy Method (MEM) analysis

Lifetimes and amplitudes extracted from the discrete analysis of NATA fluorescence intensity decay with a forced two-exponential model [**Table 4.1, 4.2 and 4.3**] do not reveal the true physical scenario due to difficulty in resolving closely spaced lifetimes and challenges in recovering amplitudes and lifetimes from multiexponential decays. An alternative approach to investigating the influence of charged amino acids on the fluorescence lifetime of NATA was attempted using the Maximum Entropy Method (MEM). It is based on a model-free approach to analyse the fluorescence intensity decays and generate fluorescence lifetime distributions. Since the MEM analysis is based on a model-free approach, it was not necessary to constrain for a 2-exponential lifetime; rather, we fit it for 100 exponentials. As seen in **Figure 4.9**, Arginine, which showed the most quenching with higher k_q, displayed a two-lifetime distribution where an additional shorter lifetime distribution was observed (for concentration > 200 mM), although the two-lifetime distribution for Aspartate and Histidine was not evident. This lifetime was similar to those originating from Arginine alone when excited at 295 nm (τ < 1ns). In contrast to the observations of MEM analysis, the discrete analysis method could not reveal the distribution of the short lifetime components. The fitted residuals for the MEM analysis, as shown in **Figure 4.10,** showed a similar pattern as that of the residuals for 2-exponential fit in the NLLS method. Thus, MEM lifetime distributions distinctly brought out

the additional luminescence contribution originating from charged amino acids, which could account for deviations observed in single exponential decay fits of NATA. Although, it should be noted that the shorter lifetime distribution was only observed when the amino acids were in excess compared to NATA and that too for the amino acid with a higher quenching propensity than others.

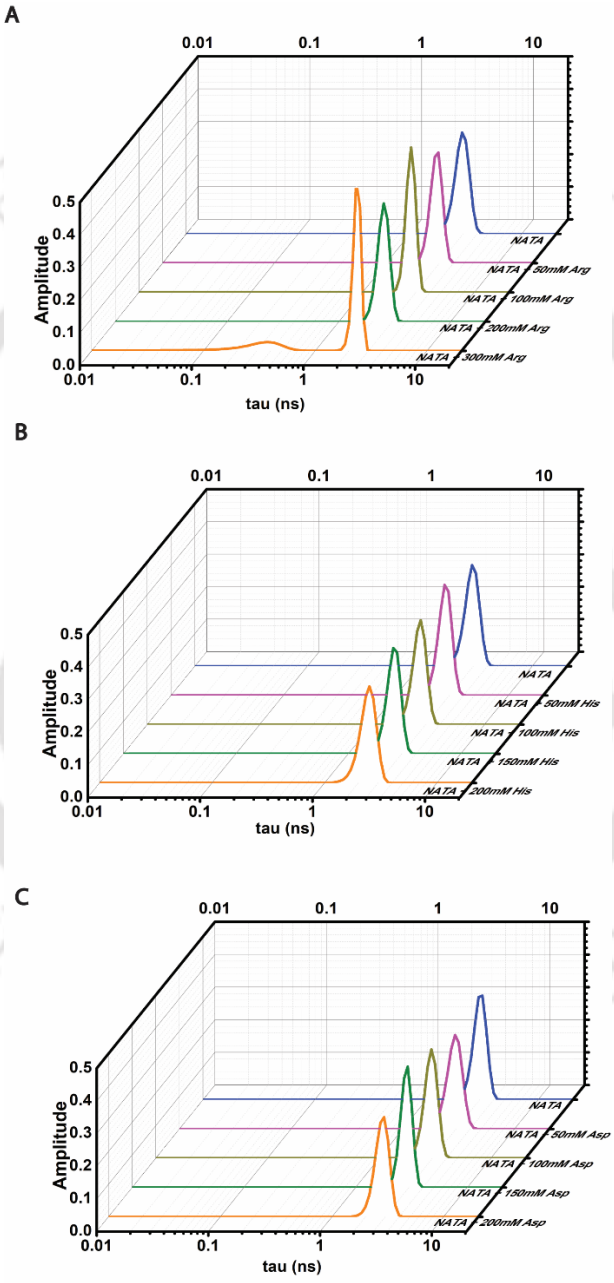


Figure 4.9: MEM Lifetime Distribution for NATA in the presence of charged amino acids. Subfigures A, B and C show the lifetime distribution for increasing concentrations of Arginine, Histidine and Aspartate, respectively, compared to NATA.

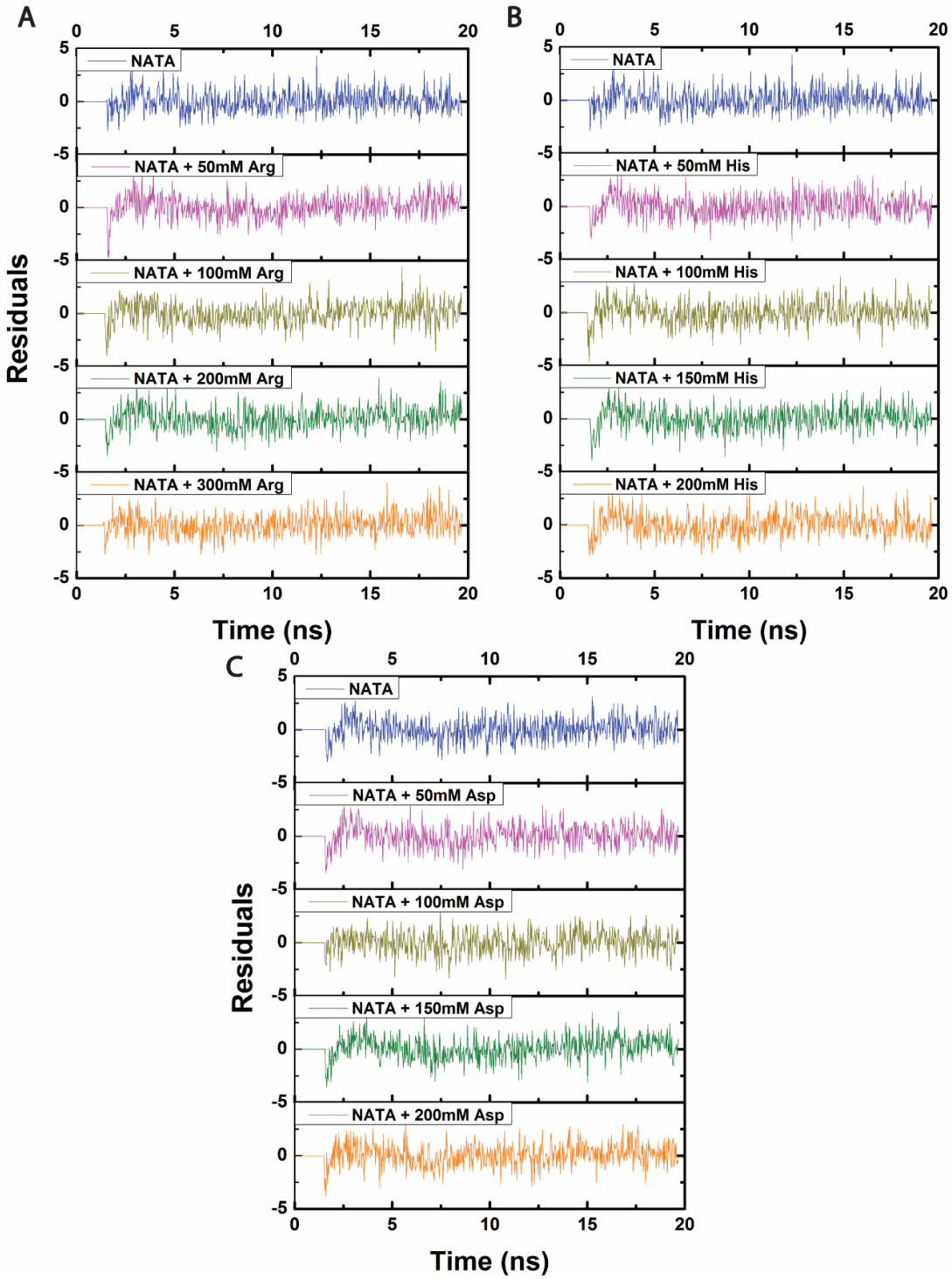


Figure 4.10: Residuals for MEM Lifetime Analysis for NATA in the presence of charged amino acids. Subfigures A, B and C show the residuals in the presence of different concentrations of Arginine, Histidine and Aspartate, respectively, compared to NATA alone.

4.2.3. Fluorescence intensity decay analysis of NATA in the presence of a monomeric protein: Symfoil PV2

4.2.3.1. Non-linear Least Square (NLLS) analysis

As discussed in the earlier section, although it could be said that the charged terminals in the amino acids contribute to the complexity in NATA fluorescence and mimic Trp-like behaviour in charge-rich proteins, it would be justifiable to validate the statement only when we explore NATA in the presence of Arginine, Aspartate and Histidine rich proteins. Thus, we repeated the experiments and replaced the amino acids with Symfoil PV2, a monomeric protein with no aromatic amino acids with a high content of Arginine, Aspartate and Histidine residues. But before that, we checked the steady-state fluorescence anisotropy of NATA in the presence of Symfoil PV2 to confirm whether NATA was in close proximity to the protein. To our surprise, in **Figure 4.11**, we found that the anisotropy did change marginally, suggesting some non-specific binding to the protein. This observation, in a way, favoured our study as it would behave like having a Tryptophan placed inside a protein, and the proximity of charged amino acids side chains in the protein around the indole ring would create a more realistic environment. Further compared to amino acids, far less concentration of protein is needed to create ambiance of charged amino acids The fluorescent intensity decay of NATA with varied concentrations of Symfoil PV2 was collected in the 320-360 nm range with a band pass filter and analysed when excited at 295 nm.

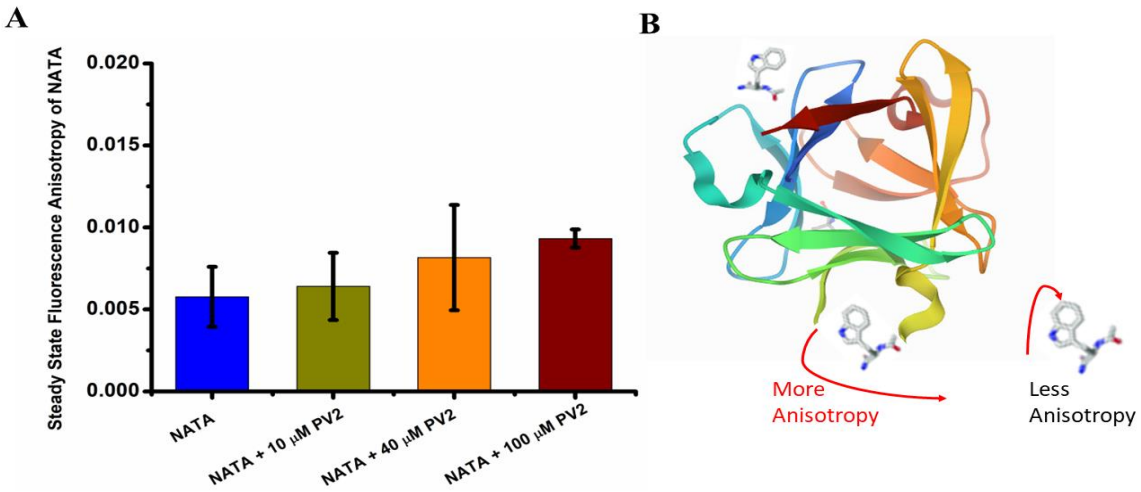


Figure 4.11: Steady-state fluorescence Anisotropy of NATA in the presence of Symfoil PV2. Subfigure A shows the anisotropy of NATA in the increasing presence of Symfoil PV2. Subfigure B shows a schematic representation of the non-specific binding of NATA to Symfoil PV2.

Here, the 1 and 2 exponential models were used to fit the fluorescent intensity decay, as shown in **Figure 4.12**. It has been found that the two-exponential model fit much better for all the samples than its one exponential model as the reduced Chi-Square, χ^2_R values of close to 1 depicted a better fit.

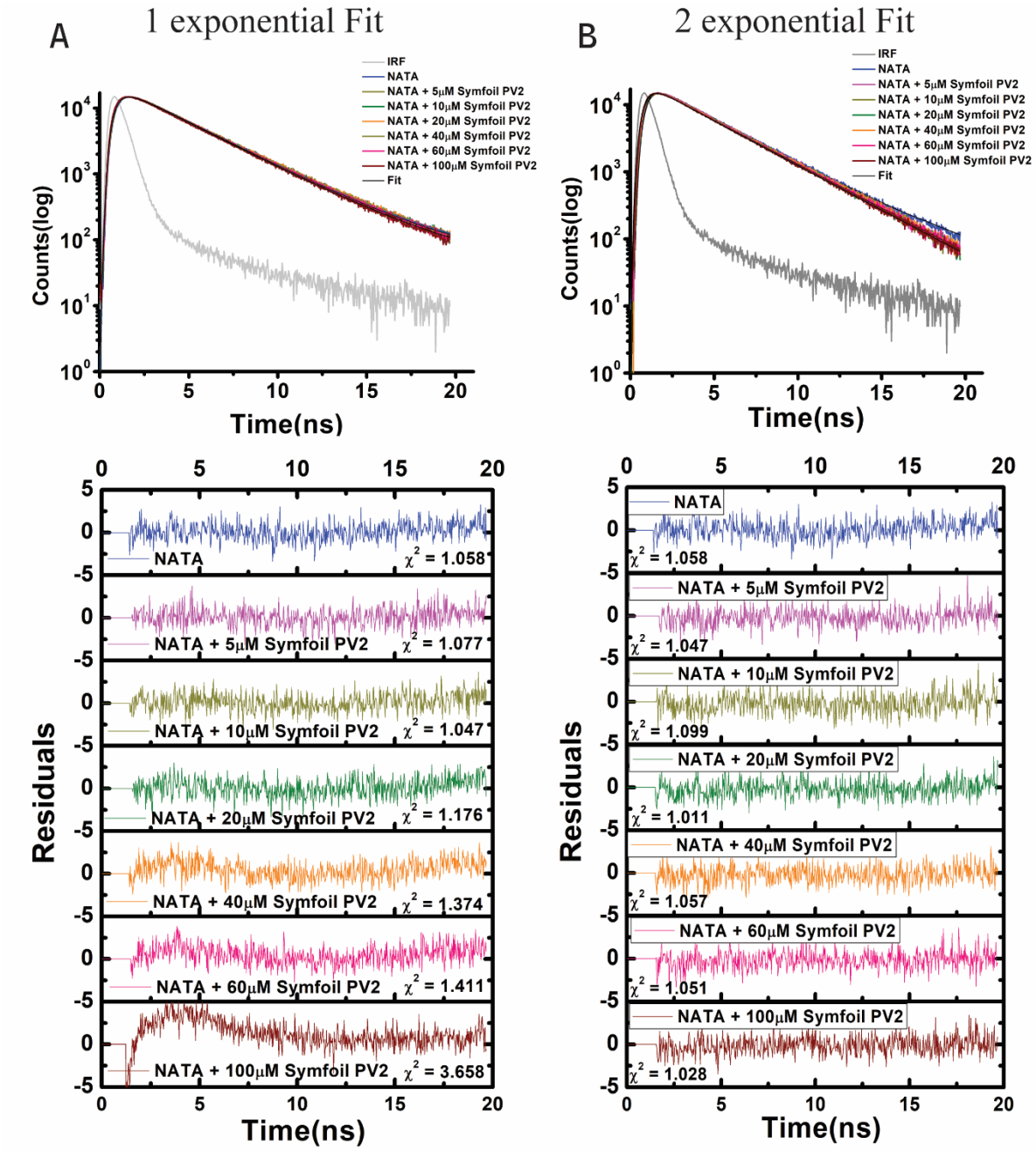


Figure 4.12: Fluorescence lifetime decay analysis of NATA in the presence of Symfoil PV2 at 295 nm. Subfigures A and B show the fitted decays for one exponential and two-exponential fit, respectively, along with their residuals.

	1 exponential-fit		2 exponential-fit					
	τ_{mean} (ns)	χ_R^2	α_1	α_2	τ_1 (ns)	τ_2 (ns)	τ_{mean} (ns)	χ_R^2
NATA	3.10 (0.003)	1.058	1	0	3.10 (0.003)		3.10 (0.003)	1.058
NATA+5 μM PV2	3.08 (0.033)	1.077	0.864	0.136	3.33 (0.031)	1.44 (0.131)	3.06 (0.031)	1.047
NATA+10 μM PV2	3.08 (0.002)	1.047	0.875	0.125	3.33 (0.036)	1.51 (0.212)	3.07 (0.012)	1.099
NATA+20 μM PV2	3.09 (0.007)	1.176	0.853	0.147	3.34 (0.035)	1.52 (0.012)	3.07 (0.023)	1.011
NATA+40 μM PV2	3.07 (0.004)	1.374	0.843	0.157	3.366 (0.017)	1.57 (0.044)	3.08 (0.012)	1.057
NATA+60 μM PV2	3.04 (0.007)	1.411	0.829	0.171	3.33 (0.017)	1.65 (0.006)	3.04 (0.021)	1.051
NATA+100 μM PV2	3.01 (0.010)	3.658	0.799	0.201	3.32 (0.004)	1.62 (0.048)	2.98 (0.007)	1.028

Table 4.4: Parameter values of fitted decay for NATA in the presence of Symfoil PV2 at 295 nm excitation for 1 and 2-exponential fit. τ and α denote the lifetime and amplitude obtained from the fitted data, respectively.

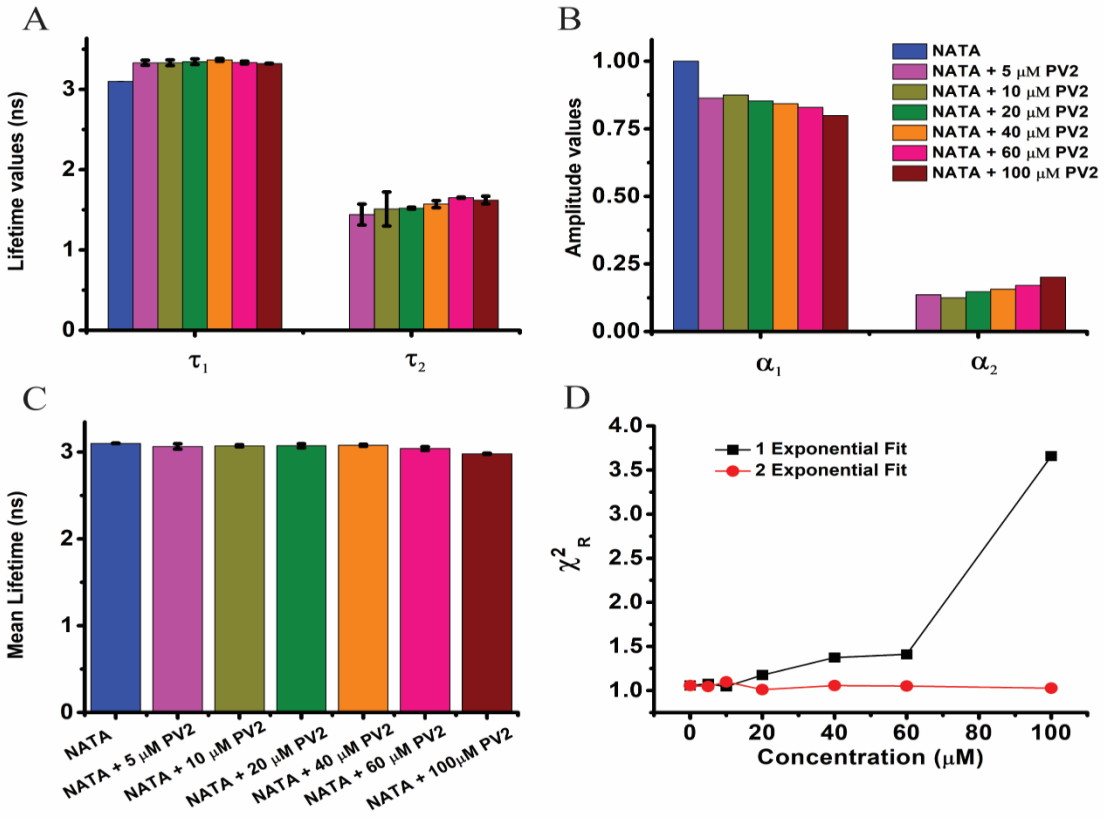


Figure 4.13: Subfigures A, B and C show the individual fluorescence lifetimes, respective amplitudes and mean lifetimes, respectively, from NLLS analysis of NATA in the presence of Symfoil PV2 at 295 nm. Subfigure D shows the reduced Chi-Square obtained from the 1 and 2 exponential fits.

It could be seen in **Figure 4.13D** and **Table 4.4**, although for lower protein concentration (< 20 μM), the decay could be fitted in one exponential, but with higher concentrations of protein, the decay fitted better in the 2-exponential model. The χ^2_R value even reached around 3.658 in one exponential fit for 100 μM concentration, whereas, in the case of the two-exponential fit, the χ^2_R values were hardly above 1.06. The better randomness in the residuals of the 2-exponential fit compared to the 1-exponential fit, seen in **Figure 4.12,** also justified the result. On close inspection, we could notice that although the concentration of protein used was very low (in order of μM) compared to amino acids (in mM), the deviation from one exponential was surprisingly far more significant in proteins (χ^2_R of 3.658) than that of amino acids (χ^2_R of 1.302 for Arginine). It thus suggested that the contribution to the complex lifetime distribution in NATA fluorescence is far more dominant in the case of proteins as compared to amino acids. Strikingly, the decrease in the mean lifetime was smaller in the case of protein than amino acids as shown in **Figure 4.13C**. However, the deviation in χ^2_R values was in contrast to the changes in mean lifetime values. It could be because the NATA fluorescence was marginally quenched in the case of proteins compared to amino acids. Moreover, the emergence of the shorter fluorescence lifetime with increasing protein concentration could be observed in **Figures 4.13A** and **4.13B,** with a marginal increase in their amplitude values, with no change in the individual lifetime values.

4.2.3.2. Maximum Entropy Method (MEM) analysis

We performed the Maximum Entropy Method to further understand the contribution of the shorter lifetime component and its distribution as analysed from the 2-exponential fit. Good randomness in the residuals was observed in the MEM analysis, as seen in **Figure 4.15**. We could clearly see in **Figure 4.14** that a shorter fluorescence lifetime emerged with the increase in protein concentration (>60 μM), which was earlier absent in NATA alone. Moreover, unlike amino acids, the longer lifetime did not shift towards lower values, and only the contribution from each component changed. Thus, this result is in agreement with our Discrete analysis data (**Figures 4.13 A** and **B**), where the amplitude values of each lifetime component (α_1 and α_2) changed with no change in the individual lifetimes (τ_1 and τ_2), with a marginal decrease in mean lifetimes signifying no possible quenching.

Our observations thus indicate that a charged protein is more potent towards contributing additionally to the fluorescence lifetime in NATA than charged amino acids, resembling Trp fluorescence. Even when taken in excess, free amino acids cannot create a 3D localised charged environment around NATA. In contrast, in proteins, the close spatial 3D proximity of variously charged termini around NATA makes a more confined charge cluster, thus influencing Trp emission with stronger luminescence.

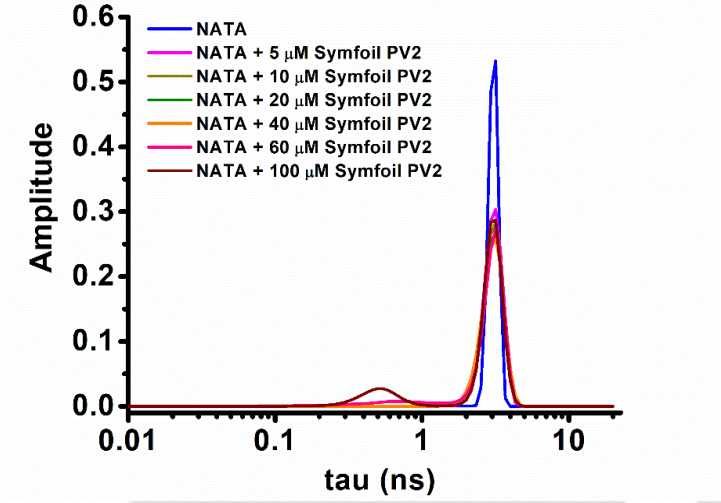


Figure 4.14: Fluorescence Lifetime distribution of NATA in the presence of Symfoil PV2 from MEM.

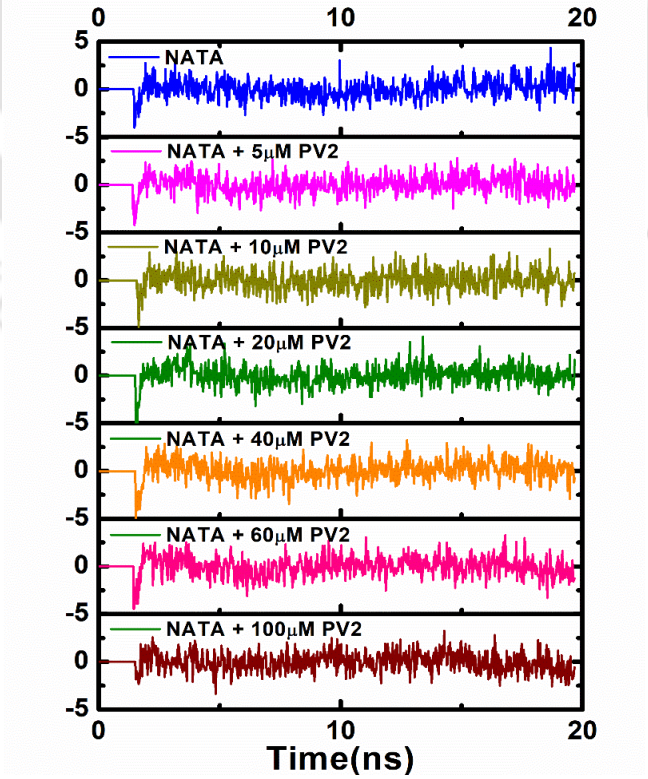


Figure 4.15: Residuals for MEM analysis of NATA in the presence of increasing concentrations of Symfoil PV2.

4.3. Conclusions

- This study brings forward the contribution of charged amino acids, namely Arginine, Aspartate, and Histidine, which can drastically influence Trp fluorescence. The luminescence from the amino acids alone mixes with the fluorescence of NATA or indole, thus mimicking the Trp-like fluorescence as observed in proteins.

- Both the Discrete and MEM analysis shed light on the phenomenon underlying the origin and contribution of the shorter fluorescence component from the indole ring in Trp.

- One exciting observation is the propensity of these charged residues to sully the indole fluorescence, irrespective of whether its fluorescence is quenched or not.

- Amino acids in clusters quench the NATA fluorescence more than that of proteins, but the close 3D proximity of charged clusters in protein is more potent in contributing to the complexity of the fluorescent lifetime of the indole group.

- Although, ProCharTS luminescence cannot compete with Trp fluorescence due to lower luminescence output, the possible interference from the novel ProCharTS luminescence from the side chain of such non-aromatic charged residues should be taken into account.

Chapter 5

To monitor the early events of Amyloid beta (Aβ) Switch peptide aggregation in real-time using ProCharTS

5.1. Introduction

Alzheimer's disease (AD) is a life-threatening neurodegenerative disorder related to the Aβ 40 or Aβ 42 peptide, where insoluble amyloid fibrils deposit in the neuropil. The initial fibrillary oligomers of Aβ are thought to be the potential reason behind it. Thus, inhibiting the peptide oligomer and protofibrils formation in its early stages by identifying potential drugs is important for treating AD. Here, in this study, the early stages of aggregation of the hydrophobic core segment $A\beta_{16-22}$ (KLVFFAE) are being analyzed using switch peptides (SwP). The pH-dependent O->N acyl migration in the molecular switches inserted deliberately in the $A\beta_{16-22}$ sequences enables the systematic study of conversion of random coiled peptides into β-sheets. The pH change from acidic (~pH 2) to neutral or slightly basic (~pH 6.8-7.2) dictates the acyl migration.

Limitations from various expensive techniques prompted us to look for simpler, cost-effective, non-invasive techniques to monitor the aggregation of Aβ peptides. Here, we used intrinsic absorbance/ luminescence arising from ProCharTS to address the issues often experienced by other techniques. ProCharTS signal is expected to increase due to the increased propensity of the proximal contacts between charged terminals of Lys/Glu residues in each anti-parallel β-strand having distances less than 4.8 Å[91]. It eliminates the higher-end, expensive techniques that do not provide information on molecular level interactions but only on the gross shape and size. The non-invasive nature of ProCharTS minimizes the effect often experienced by extrinsic dyes during aggregation, which can alter the kinetics. Furthermore, the recent report on the photochemical features of intrinsic luminescence in monomeric proteins and its applicability to screen drugs for early inhibition of amyloids[176] encourages the use of ProCharTS in peptide aggregation.

Below mentioned are the different peptide cases which have been used in this study.

Pep I: Ac-SLKLVFF(H^+) SELSG-$CONH_2$ (50 μM)

Pep II: Ac-SLKLVFF(H^+) SSLSG-$CONH_2$ (50 μM)

Pep III: NH_3^+-SLKLVFF(H^+) SELSG-$CONH_2$ (200 μM)

Pep IV: NH_3^+-SLKLVFF(H^+) SSLSG-$CONH_2$ (200 μM)

Pep V: Ac-SLKLVFF(H^+) SELSG-$CONH_2$ (200 μM)

5.2. Results and Discussions

5.2.1. Circular Dichroism (CD) to detect secondary structure changes in SwP aggregation

CD was performed to assess the time frame of the early events of the aggregation of SwP during which the secondary structural transitions took place. Based on this time scale, all the other related techniques were carried out for the same peptides. Aggregation of SwPs started when

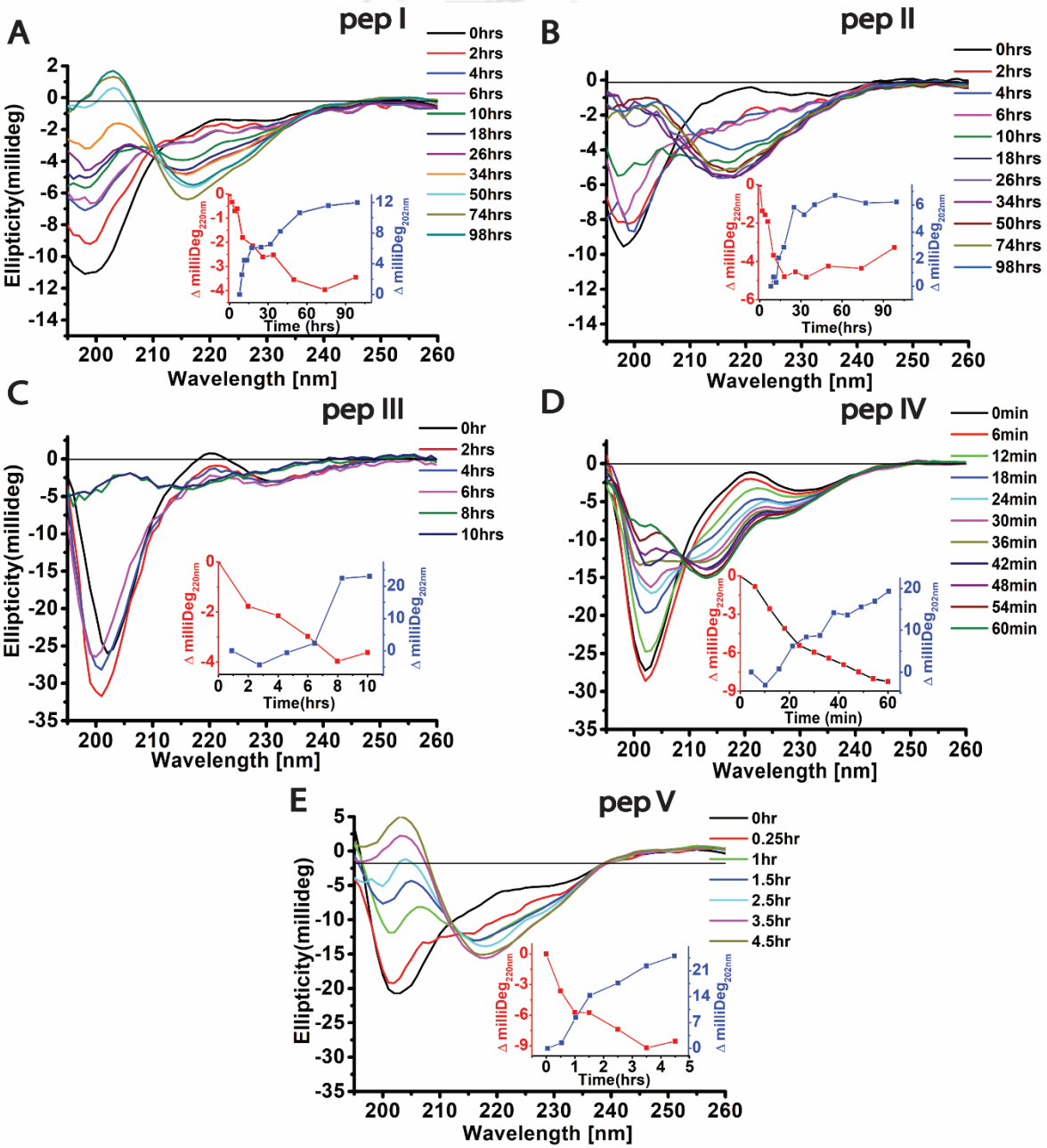


Figure 5.1: CD spectra for switch peptide aggregation at different times in aggregation buffer (phosphate buffer, 50 mM, pH 7.2). Inset shows the change in ellipticity at 220 nm and 202 nm versus time.

the peptides were transferred in pH 7.2 (50 mM phosphate buffer). Different time frames were observed for each peptide, where a complete secondary structural transition from random coil to β-sheet structure takes place. As shown in **Figure 5.1**, all the switch peptides had an ellipticity value with minima near 190-200 nm, characteristic of a random coil. As the aggregation proceeded, the ellipticity at 190-200 nm increased, with a decrease at 220 nm, which characterizes β-sheet formation. The changes in the ellipticity at 220 nm and 202 nm were plotted as shown in **Figure 5.1 [insets]**. For all the SwPs, a sigmoidal decrease in ellipticity at 220 nm was evident. The sigmoidal profile signifies the complete transition from random coil to β-sheet structure, as the ellipticity values flattened out. Interestingly, the ellipticity at 220 nm increased at even longer time scales. It hints that higher-order insoluble mature fibrillary structures formed precipitates as the aggregation proceeded further. We thus restricted ourselves to the study of only the formation of soluble oligomeric aggregates in the initial stages of aggregation kinetics, with minimal contribution from light scattering, which arises from mature fibrils.

For all the SwPs, the change in ellipticity at 220 nm was consistent, and a gradual decrease was observed. However, in the case of pep III, an abrupt reduction in ellipticity at 190-200 nm between 6-8 hrs was observed, suggesting that pep III had a lag phase in its aggregation mechanism till 6 hrs. The fastest aggregation kinetics was observed for pep IV (~1 hr), followed by pep V (~4 hrs). The slowest kinetics were observed for pep I (~98 hrs), and pep II (~ 98hrs), where the peptide concentrations used were at 50 μM. For all other cases, 200 μM concentrations were used since with 50 μM concentration, aggregation could not be performed. Thus, the higher initial concentration of the peptide increases the seeding of peptides in aggregation kinetics, speeding up the process. Our results are consistent with similar other reports for Aβ aggregates[149,177]. A similar timeframe was used to study the Aβ peptides oligomer formation and subsequent fibril formation using conventional fluorescence-based extrinsic probes: ANS and ThT, which is discussed in the next section.

5.2.2. 8-Anilinonaphthalene-1-sulfonic acid (ANS) and Thioflavin T (ThT) assays

8-Anilinonaphthalene-1-sulfonic acid (ANS) is a fluorescent probe commonly used to assess the hydrophobic content in proteins. On excitation at 380 nm, it shows a manifold increase in fluorescence intensity at 480 nm when bound to hydrophobic cores of proteins or assembly, as compared to being free in solution with minimal fluorescence at 520 nm. Moreover, the change

was characterized by a blue shift with an increased fluorescence signal. Similarly, when excited at 450 nm, an increased fluorescence at 480 nm in ThT-bound fibrils was observed as compared to free ThT, with a slight blue shift from 490 nm to 480 nm in aggregates with increased fluorescence[47]. Here, the ANS assay is expected to give more information about the early events of the SwP aggregation where the peptides come nearby and form hydrophobic pockets, compared to ThT.

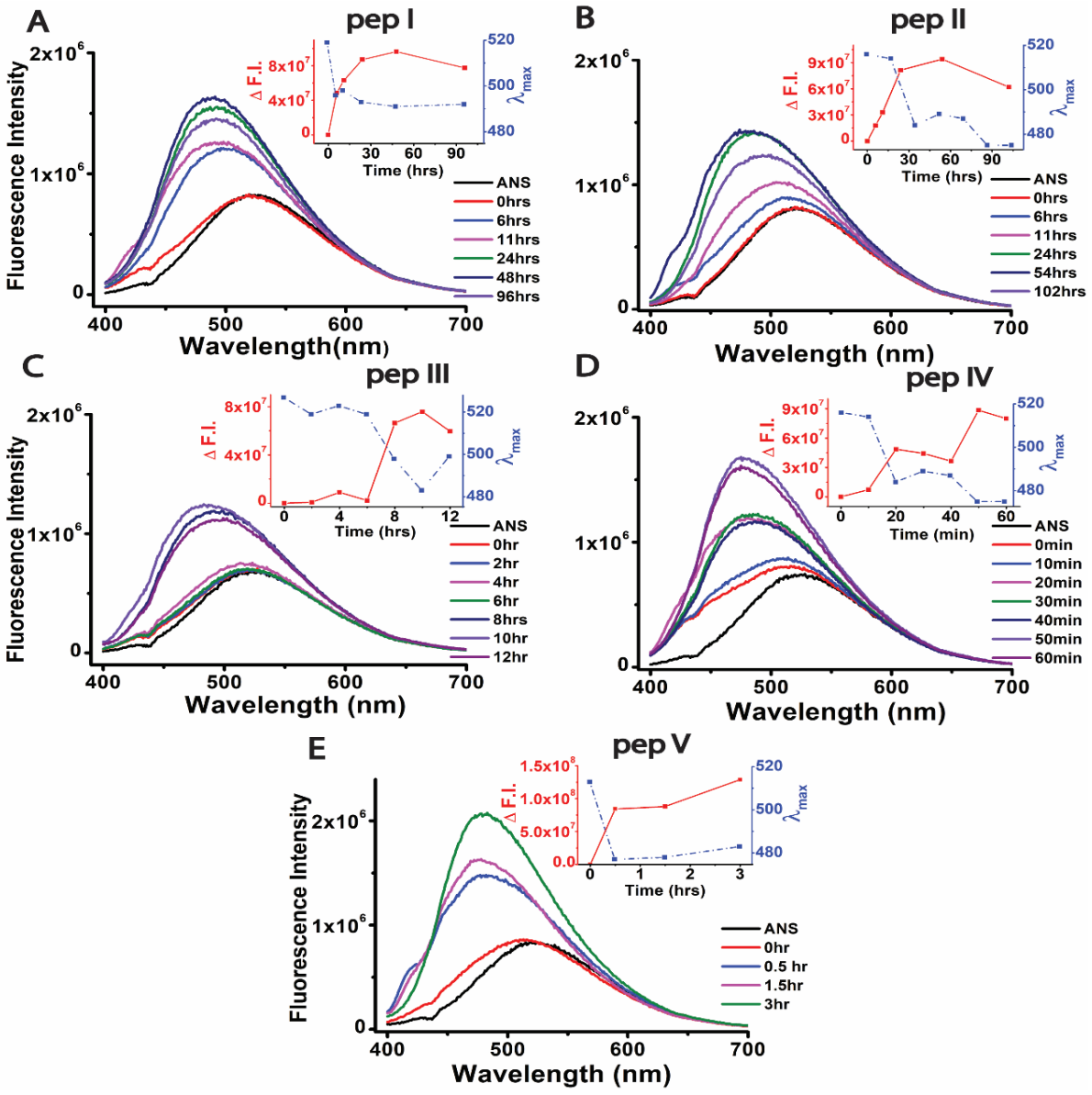


Figure 5.2: ANS spectra for switch peptide aggregation in phosphate buffer, 50 mM, pH 7.2. The concentration of ANS was kept at 20 μM with peptides in a 1:1 ratio. All the readings were done in duplicates with 2 scans each. Inset shows the change with respect to 0 hrs in integrated fluorescence area and the shift in emission maxima versus time.

Similar to CD analysis, the aggregation of all the switch peptides was recorded in similar conditions. The data acquisition was made at different time intervals keeping the maximum time duration of the aggregation identical to that of the CD to compare the aggregation kinetics with that of ANS and ThT.

As can be seen in **Figure 5.2**, all the SwPs, upon aggregation, showed 2 to 3-fold increase in ANS fluorescence with prominent blue-shifted emission maxima (λ_{max}). Peptides in monomeric states showed λ_{max} at ~520 nm, while for the aggregated forms, the λ_{max} shifted to ~480 nm. The integrated fluorescence area under the curve of each spectrum was calculated and shown in the respective insets. All the peptides showed a gradual sigmodal change in the integrated ANS fluorescence intensity. The lag phase observed in the CD analysis in Pep III where the ellipticity at ~202 nm increased quite drastically, was also evident in the ANS assay, as shown in **Figure 5.2C**. Moreover, the decrease in the λ_{max} profile for all the peptides also complemented the increasing fluorescence intensity pattern. The respective time frames for the assay completion seemed to match quite well with CD analysis, which will be discussed in ***Section 5.2.7***. From the ANS study, it can be said that ANS could give us similar insights as compared to CD, where the formation of hydrophobic clusters during the conversion of the random coil to β-sheet is observed.

Similarly, ThT assay picked up similar kinetics compared with CD and ANS data, as shown in **Figure 5.3.** Increased fluorescence intensity at ~487 nm was observed with time in pep I, pep II, pep V. A 4 to 5-fold increase in the ThT fluorescence upon aggregation was observed, suggesting aggregate formation in the SwPs. However, as seen in **Figures 5.3C** and **5.3D**, not much change was observed for pep III and IV. One exciting observation was that both these peptides were non-acetylated, whereas the other three peptides, which showed increased fluorescence, were acetylated. The reason for this marginal increase is not apparent at this point. However, the binding of ThT being hampered by the presence of the free N-terminal amino group can be a possible reason behind the observation. Similar reports of ThT altering the kinetics or perturbing the assembly mechanism have been observed[178]. Moreover, when the change in ThT intensity was plotted versus time for the 3 acetylated peptides, a similar sigmoidal pattern could be seen in all of them. Thus, ThT could track the aggregation kinetics but not in all cases. CD and ANS proved to be better techniques for characterizing the early events of aggregation than ThT. Optimizing the peptide to dye ratio while using fluorescent dyes like ThT or ANS pose several challenges in studying the aggregation of proteins.

Variations or erroneous data due to self-aggregation of dyes or limited dye per peptide present in the solution can hamper the aggregation kinetics[50]. Thus, we turned our attention toward more straightforward methods where the dependence on extrinsic probes can be avoided. Therefore, the aggregation kinetics of Aβ was carried out by ProCharTS absorbance and luminescence, subsequently.

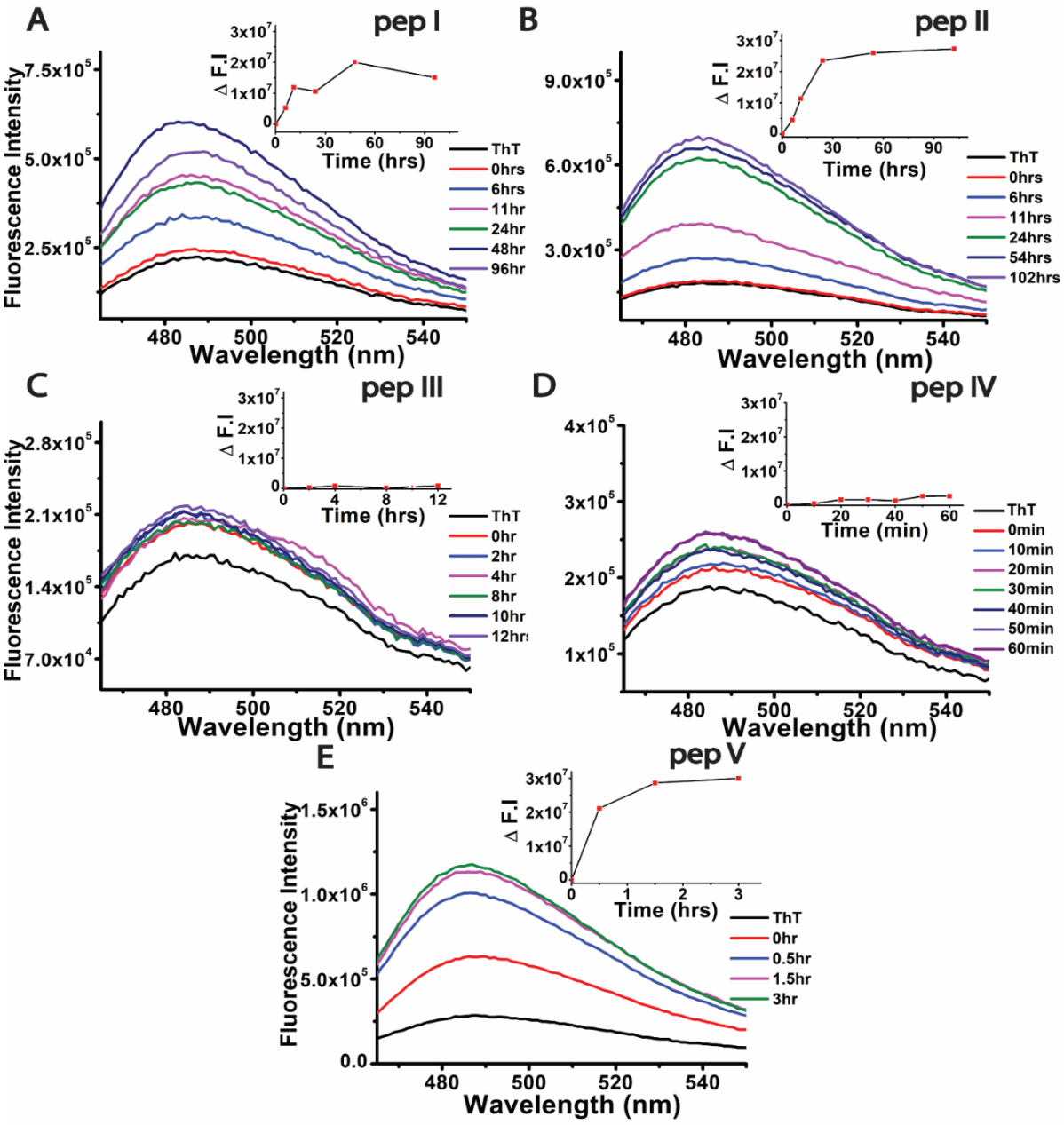


Figure 5.3: ThT assay for switch peptide aggregation at different times in aggregation buffer (phosphate buffer, 50 mM, pH 7.2). The concentration of ThT was kept at 20 μM with peptides in a 1:1 ratio. Inset shows the change with respect to 0 hrs in fluorescence integrate area versus time.

5.2.3. ProCharTS absorbance

ProCharTS absorbance of SwPs was performed in conditions similar to that of CD. During early aggregation events, the peptides stack together by the hydrophobic interactions of Phenylalanine (Phe) with multiple Phe-Phe contacts in the peptide core[179,180]. As a result, the charged residues in the flanking ends of the peptides are expected to come closer, and enhanced ProCharTS is expected in aggregates as compared to monomers. Previously, the absorbance at 220 nm and 257 nm was used to track the early events of aggregation in Aβ[149]. The absorbance at 220 nm is characterized by the amide group in the peptide backbone. In comparison, the changes in absorption at 257 nm are presumed to be caused by the π-π* transitional perturbation involved in the Phe residues during aggregation. But here, we are interested in the longer wavelength region in the UV-Visible spectrum where ProCharTS primarily dominates the absorbance.

In **Figure 5.4**, the ProCharTS absorbance from 250-800 nm is shown, whereas the spectra from 200-250 nm, are shown in the insets. As all the SwPs used in this study have only Phe as their sole aromatic amino acids, thus we observed the characteristic triple humps near 257 nm of Phe. As the aggregation proceeds, the ProCharTS absorbance in the 250-800 nm increased considerably throughout the spectrum, and the increment was consistent in all cases. Moreover, apart from the Phe peaks, no additional peaks were observed, but a prolonged absorbance tail was prominent, characteristic of ProCharTS absorption. The change in ProCharTS absorption was calculated for various definite wavelengths: 257, 295, 340, 370, 450, 600, and 750 nm. As shown in **Figure 5.4. (B, D, F, H, J)**, the increase followed a similar sigmoidal trajectory. Although the rise in ProCharTS was relatively consistent for all the wavelengths measured way till 700 nm, the sensitivity was drastically reduced at longer wavelengths. However, it should be noted that some minor contributions from scattering could sully the absorbance measured during aggregation. So, to eliminate the possibility of scattering, all the absorbance values were scatter corrected starting from 450-600 nm. The logarithmic dependence of ProCharTS absorbance and wavelength (log A vs. log λ) was used[150], and scattering for lower wavelengths was extrapolated way till 250 nm, shown in **Figure 5.5**. Thus, whatever increase in absorbance is observed is likely due to the sole contribution from ProCharTS, nullifying any possible scattering.

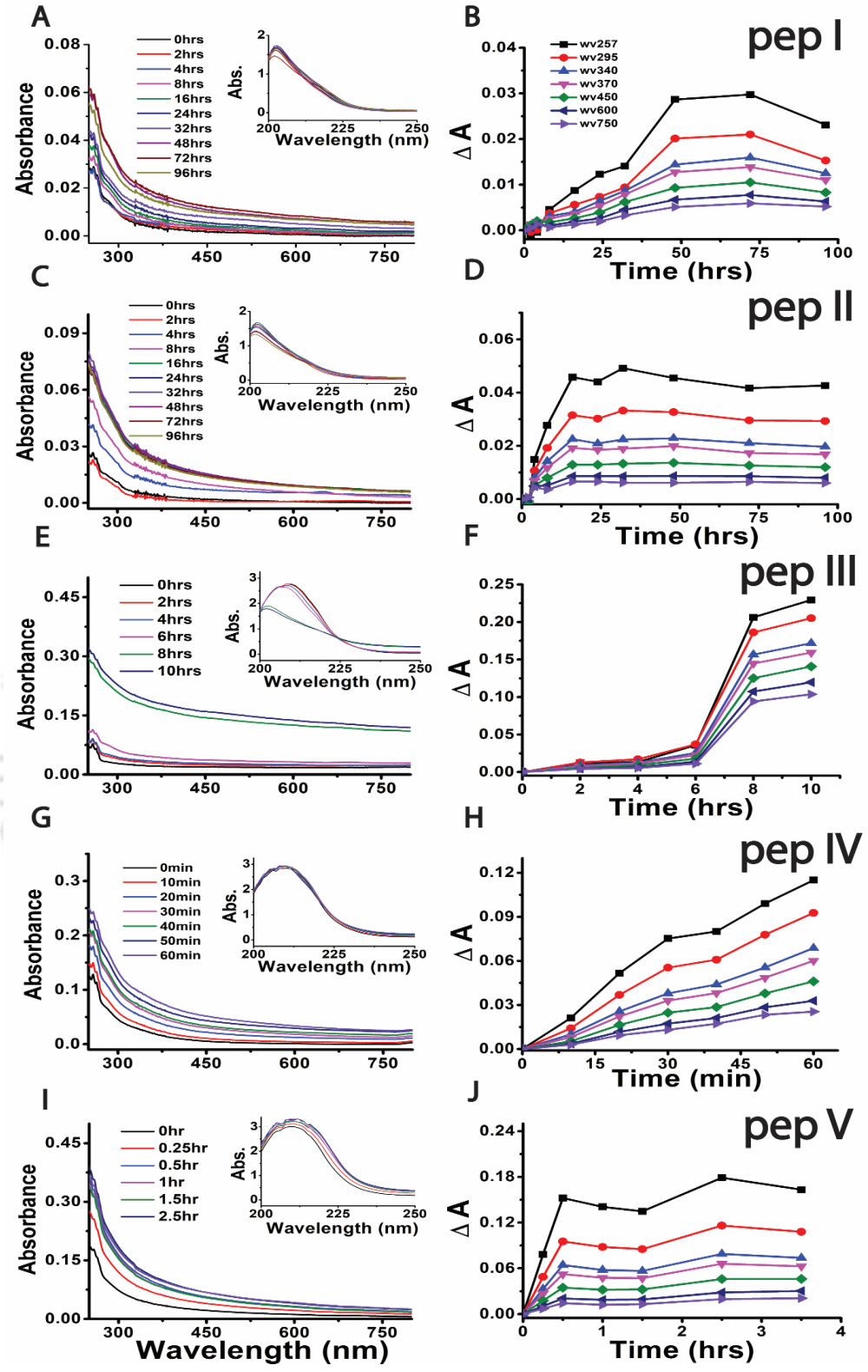


Figure 5.4: ProCharTS spectra (250-800 nm) for switch peptide aggregation at different times in aggregation buffer (phosphate buffer, 50 mM, pH 7.2) is shown in panel A, C, E, G, and I. All the readings were done in duplicate. Inset shows the identical spectra from 200-250 nm. Subfigures B, D, F, H, and J indicate the change with respect to 0 hrs in ProCharTS absorbance at various wavelengths versus time.

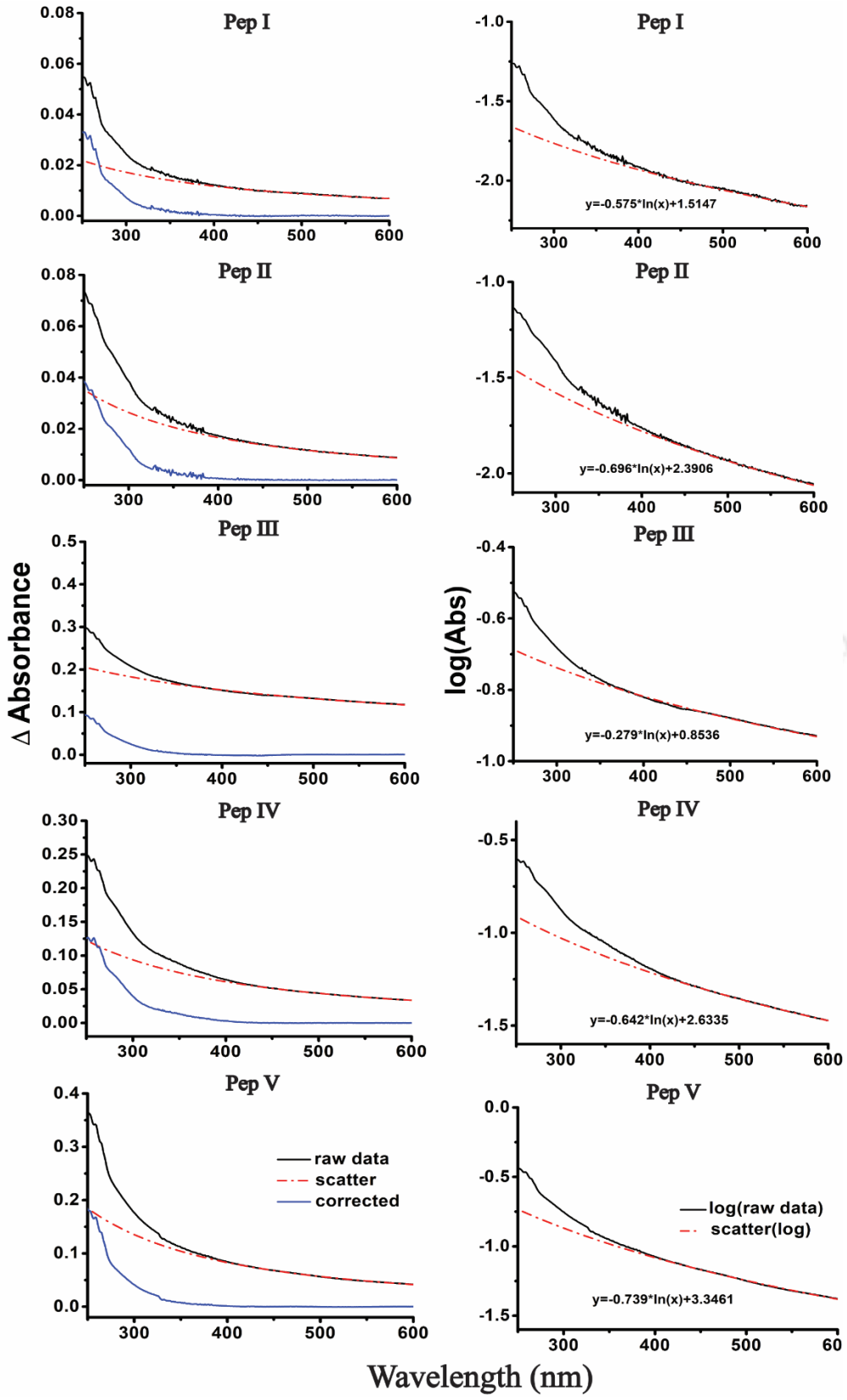


Figure 5.5: Corrected Absorbance for any possible scattering for all the SwPs are shown in left panel. The right panel shows the log A vs log λ for the scatter correction from 450-600 nm which extrapolation till 250 nm.

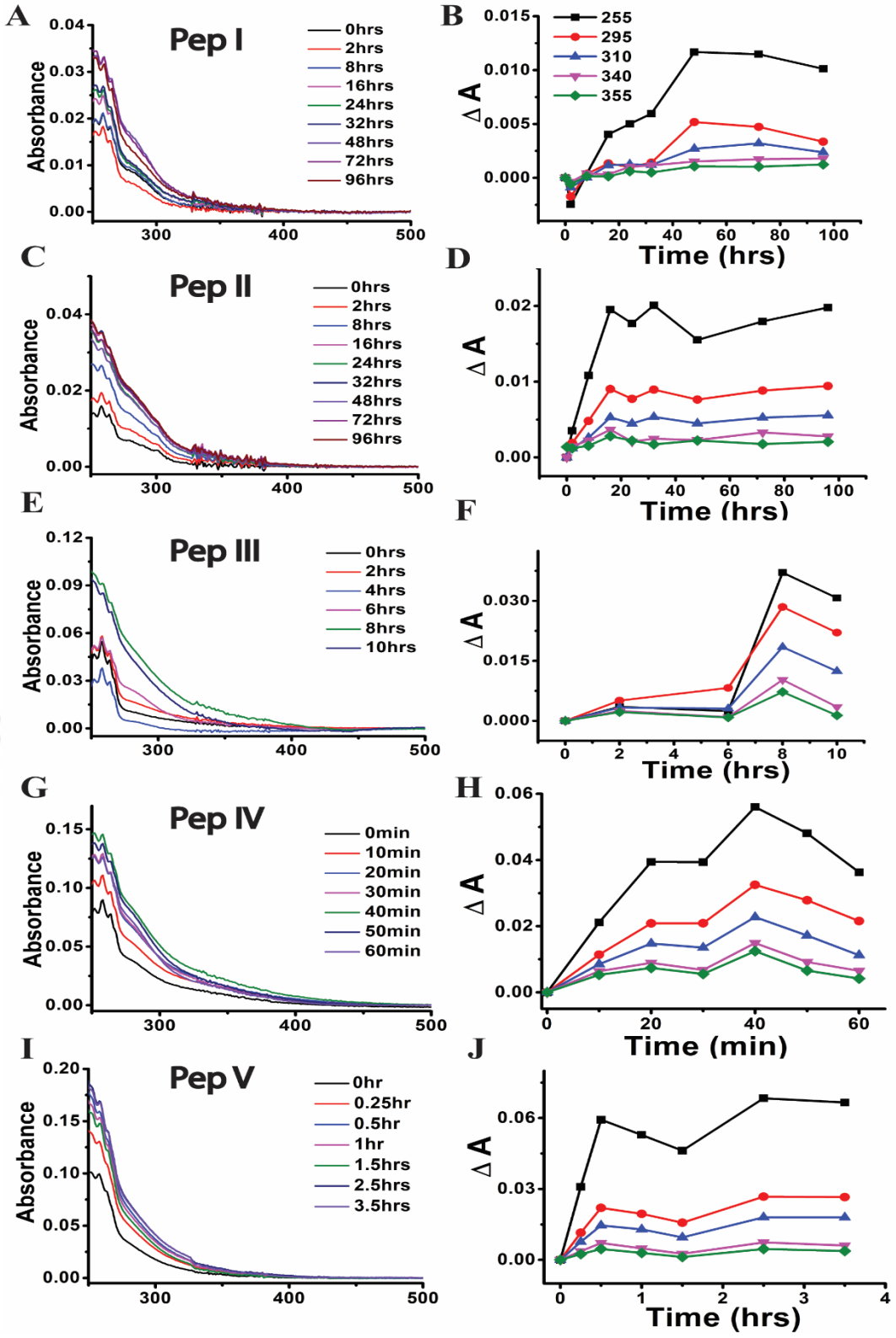


Figure 5.6: Corrected ProCharTS spectra (250-800 nm) for switch peptide aggregation at different times are shown in subfigures A, C, E, G, and I. In subfigures, B, D, F, H, and J, the change with respect to 0 hrs in ProCharTS absorbance at various wavelengths versus time is plotted.

ProCharTS absorbance after scatter-correction was plotted as shown in **Figure 5.6.** Similar to our observation before scattering correction, it too showed a consistent increase in ProChartS absorbance. Interestingly, the percentage increase in ProCharTS absorbance before and after scatter-correction was similar for most of the peptides. The ProCharTS absorbance values of each peptide were plotted with time for various definite wavelengths: 255, 295, 310, 340, and 355 nm. As seen in **Figure 5.6 (B, D, F, H, J)**, the aggregation was fastest pep IV, followed by pep V, pep III, and finally, the peptides with a lower concentration, i.e., pep II and pep I. Similar aggregation kinetics has been observed in CD, ANS, and ThT assay.

Moreover, in the case of pep III, the same lag phase was also observed as that of ANS and CD. Thus, ProCharTS absorbance is sensitive to detecting the early stages of aggregation. Similarly, computational studies on the role of charge transfer related to the structural changes in Aβ peptides and similar polypeptides also support the applicability of ProCharTS on amyloid aggregation[41,181,182]. A similar observation has been reported in di and tri peptide aggregates of Phe, with increased optical absorbance in 400-650 nm[183,184].

5.2.4. ProCharTS luminescence

Since ProCharTS absorbance has been applied successfully to characterize the early events of SwP aggregation, it prompted us to use ProCharTS luminescence to monitor the same. Moreover, both ProCharTS absorbance and luminescence are based on the same principle of electron charge transfer and subsequent recombination of the relaxed electrons. Luminescence of SwPs was collected at different times, similar to that of other techniques. Three different excitation wavelengths: 295, 340, and 370 nm, were selected, and their emission was collected between 315-570 nm, 360-660 nm, and 390-720 nm, respectively. As the aggregation continued, a gradual increase in luminescence intensity could be observed for all the SwPs, except for pep IV, as seen in **Figure 5.7**. The increment in luminescent intensity was prominent across all the wavelengths; however, the emission intensity decreased drastically with increasing excitation wavelength. Strikingly, a similar decrease in the ProCharTS absorbance with increasing wavelength was observed. One single peak for every excitation wavelength could be seen for all of the peptides, except for pep II, where one additional minor peak at ~490 nm was observed, the origin of it is unknown, and further investigation is required.

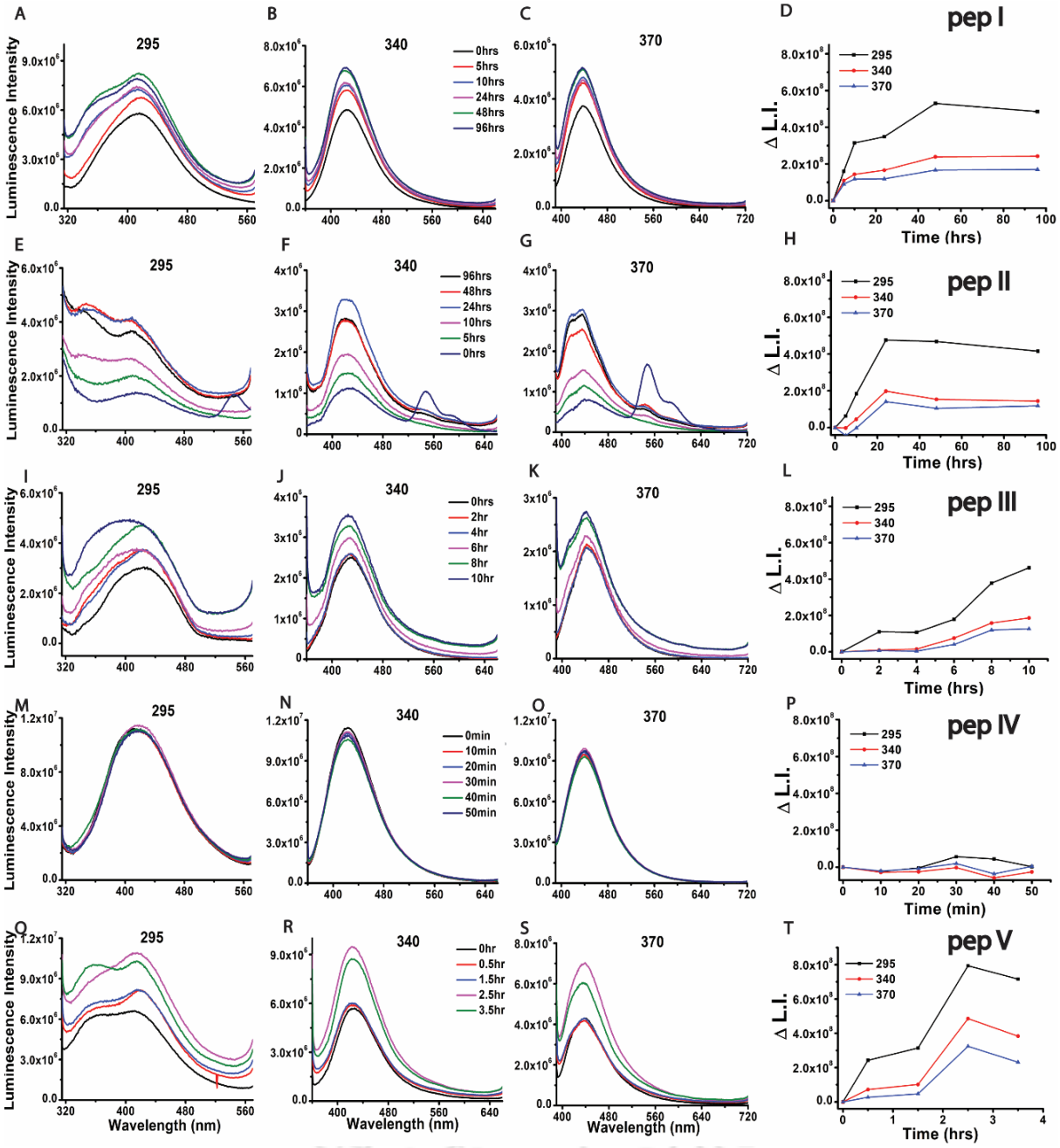


Figure 5.7: ProCharTS Luminescence Emission spectra for switch peptide aggregation at different times for 3 different excitation wavelengths of 295 nm, 340 nm and 370 nm shown in subfigures (A, E, I, M, Q); (B, F, J, N, R) and (C, G, K, O, S), respectively. Subfigures D, H, L, P, and T show the change with respect to 0 hrs in integrated luminescence versus time for individual excitation.

The integrated luminescence area under each peptide's emission curve was calculated and plotted versus time, as shown in **Figure 5.7(D, H, L, P, and T).** Similar to that of ProCharTS absorbance, the increase in ProCharTS luminescence showed a sigmoidal curve for all the peptides, except pep IV, where hardly any change could be seen, as shown in **Figure 5.7P.**

Among the peptides which showed a noticeable increase in ProCharTS luminescence, the aggregation kinetics was fastest for Pep V, followed by pep III, and then the peptides with the lower concentration, i.e., pep II and pep I. Similar kinetics was found in earlier analysis too. Interestingly, the delay observed in the case of pep III till 6 hrs was also evident from this luminescence study.

The excitation-dependent shift in emission maxima as opposed to Kasha's rule is a well-studied feature in ProCharTS luminescence often observed in monomeric proteins. During the SwP aggregation, the emission maxima (λ_{max}) shifted towards higher wavelengths with increasing excitation wavelength: emission maxima ~415 nm (λ_{ex} 295 nm); ~425 nm (λ_{ex} 340 nm), and ~440 nm (λ_{ex} 370 nm). However, irrespective of whether the peptides were in monomeric or aggregated states, the emission maxima for any particular excitation did not change as opposed to conventional chromophores like ANS and ThT, which were discussed earlier. For excitation wavelengths of 295, 340, and 370 nm, the emission maxima observed were around 415, 425, and 440 nm, respectively. It thus hints that the chromophores, although heterogeneous, are similar intrinsically and show similar traits.

Intrinsic luminescence has been recently observed in amyloids[176,185–187] in concentrated amino acids, monomeric proteins as well as oligomeric proteins and peptides[27,35]. Moreover, fluorescence waveguiding in amyloids by multiple absorption/reabsorption photon recycling has been observed[183]. Although multiple phenomena like Clustering Triggered Emission (CTE) and Aggregation Induced Emission (AIE) have been proposed, where aggregates or clusters of biomolecules are assumed to be the prerequisites for the observed luminescence. We believe that the observed luminescence intensity is due to the ProCharTS luminescence, which originates from proximal charge distribution, irrespective of protein being in monomeric or aggregated states. Aggregation only facilitates the charged residues to come closer where charge recombination of the photo-excited relaxed electron to nearby holes in the peptide backbone can occur, thereby increasing the luminescence intensity. Similar reports of charge transfer in monomeric peptides also support our hypothesis[188,189]. In summary, it can be said that ProCharTS luminescence is as good a technique as compared to other conventional methods and has the potential to be used to detect the kinetics of any aggregating proteins or peptides.

5.2.5. Time-resolved luminescence intensity decay analysis

In the earlier section, ProCharTS luminescence was used to track the early events of SwP aggregation. Here, the luminescence lifetime properties of the SwP are being studied. In contrast to steady-state luminescence, the time-dependent luminescence measures the duration an excited state electron responsible for the luminescence stays in the singlet excited state. Here, the luminescence decay features of the monomers and aggregates (at the final time point) are emphasized instead of the aggregation kinetics. Also, the relatively longer time during data acquisition for luminescence lifetime makes it impossible to track down the faster kinetics of the SwP aggregation. All the peptides were excited for this study with a 340 nm excitation pulse. The primary reason for choosing this specific wavelength was that most conventional intrinsic chromophores show absorption below 310 nm. Thus, no possible contamination from such conventional chromophores is expected to affect ProCharTS luminescence. Moreover, one of the excitation wavelengths used in steady-state ProCharTS luminescence was set to 340 nm. Thus, it seemed appropriate to perform a similar set of time-dependent luminescence studies in the same wavelength. The luminescence intensity decays for all the SwPs were analyzed by both Discrete fit and MEM analysis, as discussed below.

5.2.5.1. Non-linear Least Square (NLLS) analysis

The time-resolved luminescence lifetime decay analysis was initially done using the discrete analysis. Although ProCharTS luminescence decays for monomeric proteins have been fitted to 3 exponential fits as discussed in the previous chapter, we first tried fitting the data in both 1 and 2-exponential models. The reduced chi-square χ_R^2 for both the fits was far from unity along with non-random residuals, suggesting a bad fit. Subsequently, the 3-exponential fit showed better-reduced chi-square χ_R^2 values of less than 1.2, along with better residuals, as shown in **Figure 5.8.** The parameters obtained from the fit have been shown in **Figure 5.9** and tabulated in **Table 5.1**. As could be seen, the 3-exponential model generated three different lifetime components with varied amplitude. Whether aggregates or monomers, the shorter component (τ_1 < 1 ns) had the highest amplitude (α_1 > 0.6), followed by the intermediate lifetime component (τ_2 ~ 2-3 ns) with a relatively higher amplitude (α_2 ~0.1-0.3), while the longer component (τ_3 > 7 ns) showed the least amplitude (α_3 < 0.1). The amplitude values α_i represented the amplitude for best-fit. In contrast, the lifetime components τ_i and mean lifetime τ_{mean} values represented the averaged values for 2 or more independent measurements.

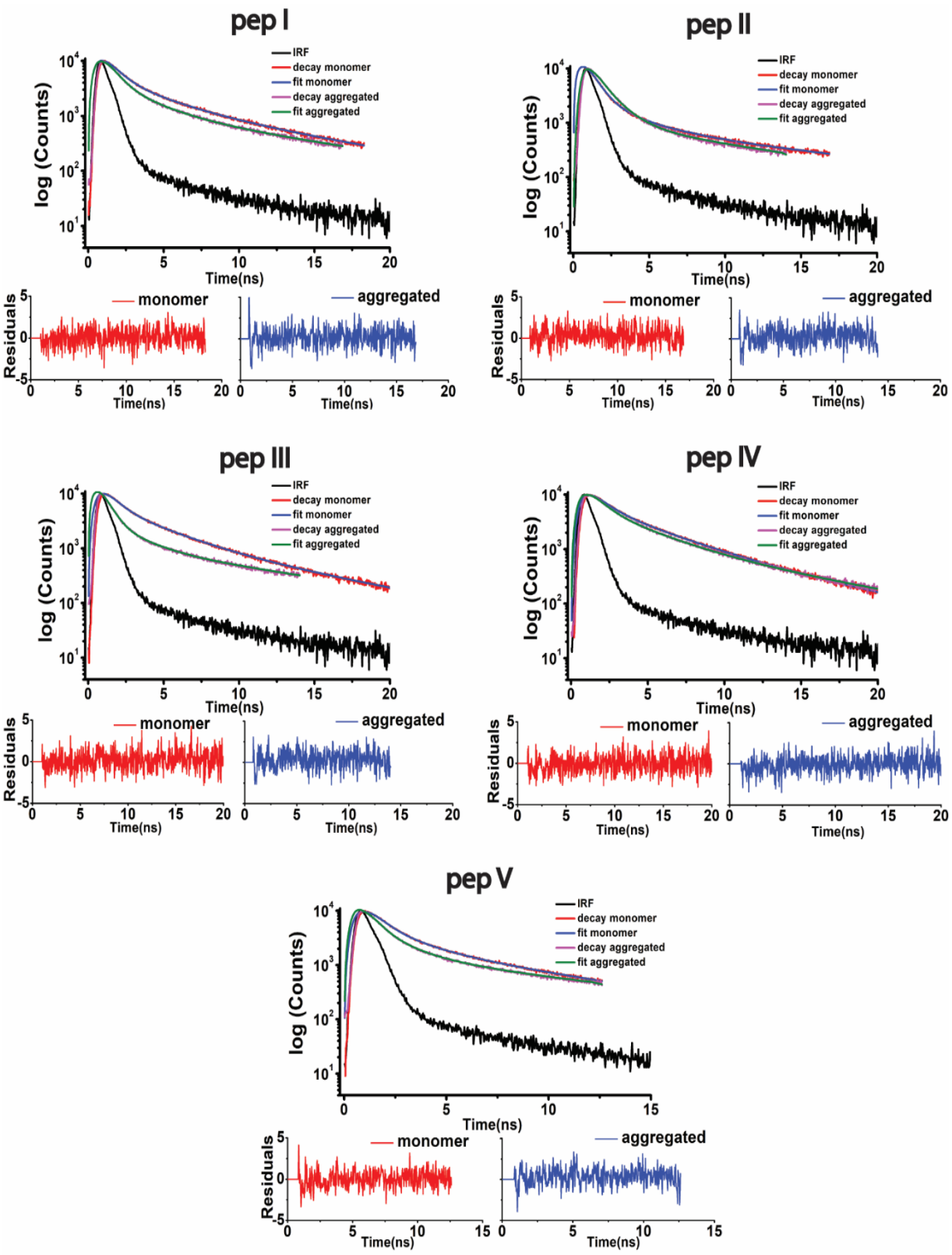


Figure 5.8: Luminescence Intensity Decay fit of all the peptides in monomeric and aggregated states when excited at 340 nm along with their respective residuals. The decays were fitted in a 3-exponential model using discrete analysis. All the measurements were made at 25 °C in duplicates, and the best-fitted data is shown here.

		α_1	τ_1(ns)	α_2	τ_2(ns)	α_3	τ_3(ns)	τ_{mean}(ns)	χ_R^2
pep I	***monomer***	0.718	0.74 (0.067)	0.196	2.5 (0.369)	0.086	9.04 (0.787)	1.92 (0.090)	1.02
	aggregated	0.787	0.70 (0.026)	0.154	2.15 (0.445)	0.059	9.74 (1.140)	1.54 (0.009)	1.18
pep II	***monomer***	0.852	0.53 (0.009)	0.118	1.90 (0.131)	0.03	13.08 (0.860)	1.08 (0.002)	1.13
	aggregated	0.639	0.30 (0.0007)	0.332	1.18 (0.074)	0.029	9.60 (0.286)	0.82 (0.030)	1.29
pep III	***monomer***	0.657	0.72 (0.033)	0.261	3.24 (0.446)	0.082	9.42 (1.042)	1.98 (0.060)	1.19
	aggregated	0.887	0.49 (0.013)	0.073	2.10 (0.139)	0.041	10.51 (0.120)	1.00 (0.020)	1.26
pep IV	***monomer***	0.596	0.85 (0.009)	0.295	3.22 (0.002)	0.109	7.34 (0.024)	2.26 (0.006)	1.12
	aggregated	0.657	0.85 (0.004)	0.269	3.302 (0.349)	0.075	11.74 (1.162)	2.08 (0.037)	1.07
pep V	***monomer***	0.708	0.57 (0.063)	0.187	2.14 (0.134)	0.105	8.76 (1.353)	1.54 (0.127)	1.00
	aggregated	0.778	0.36 (0.001)	0.176	1.38 (0.048)	0.045	9.03 (0.012)	0.94 (0.019)	1.14

Table 5.1: Parameters obtained from 3 exponential-fit for all the samples at excitation wavelength 340 nm. The individual lifetime components τ_i and mean lifetime τ_{mean} values obtained from 3-exponential fit are averaged for 2 independent measurements and shown here, while the values in the parenthesis represent their standard deviation. The amplitude values α_i along with the χ_R^2 shown here represent the best fit data from the 2 or more measurements for each sample.

As seen from **Figure 5.9C**, except for pep II, the amplitude of the shorter component in aggregates showed higher values than in monomers. While for the intermediate and the longer lifetimes, the respective amplitudes decreased marginally. Similarly, for each lifetime component, the shorter and the intermediate lifetimes had lower values in aggregates, whereas the values for longer lifetime components increased compared to the monomeric peptides. The combined effect of the variation in the amplitude and individual lifetime values between the monomers and aggregates is reflected in their mean luminescence lifetime. Peptides in the aggregated states showed a shorter mean luminescence lifetime than the monomers. The

increased contribution from the shorter lifetime components along with the shortened individual lifetime values in aggregates as compared to monomers resulted in the decrease in their mean luminescence lifetime. However, in the case of pep II, the shortened individual lifetime for all the components outweighed the decreased contribution from the shorter lifetime component, thus contributing to the overall shorter mean lifetime.

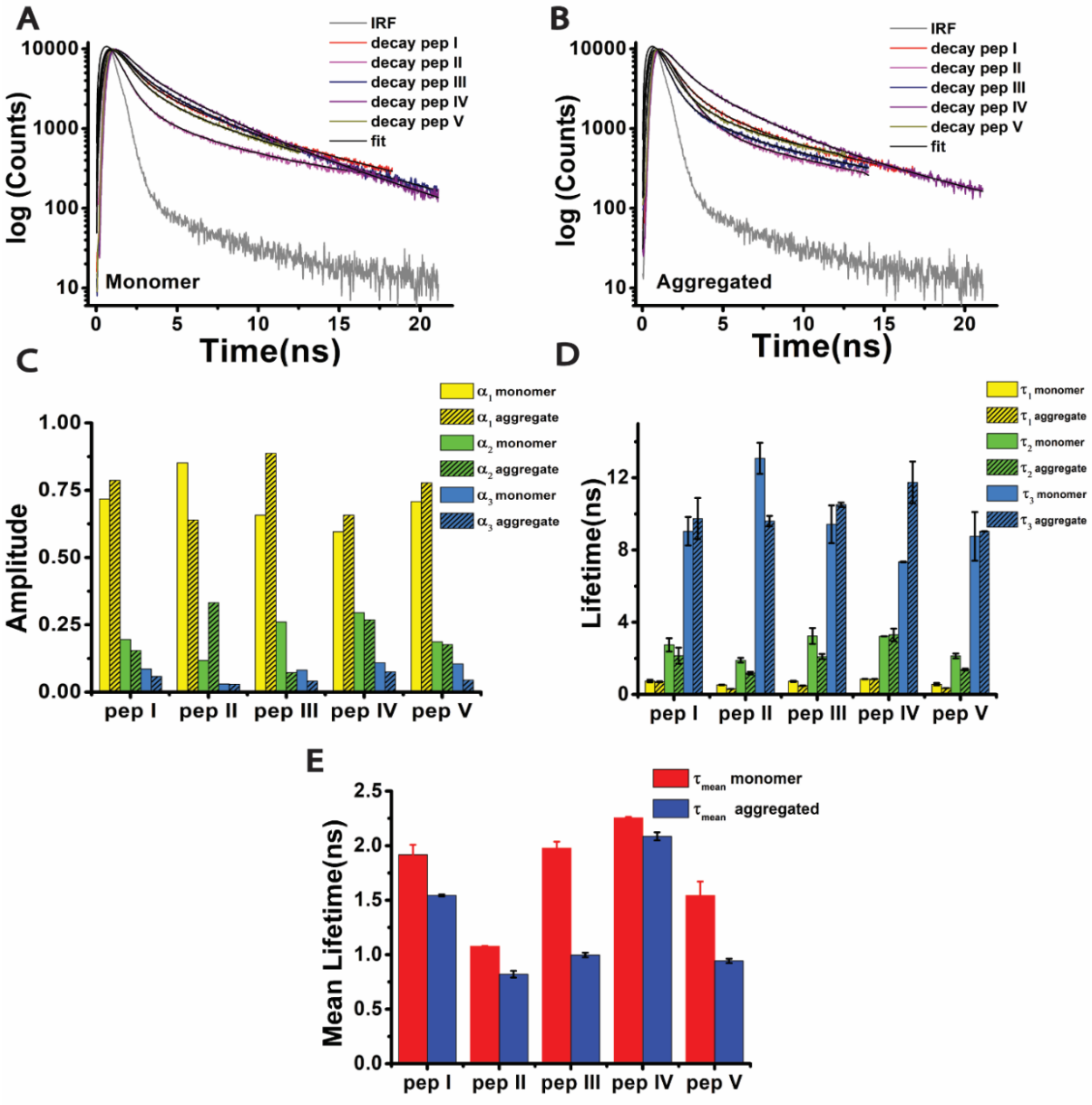


Figure 5.9: Luminescence intensity decay at 340 nm excitation along with the fitted dataset of switch peptides in the monomeric state (A) and in the aggregated state (B). Individual luminescence lifetime (D) along with their amplitude (C) and the mean lifetime τ_{mean} values (E) are indicated in the panel. The individual lifetime components τ_i and mean lifetime τ_{mean} values are averaged for 2 independent measurements. The amplitude values α_i shown here represent the best-fitted data.

The emergence of three luminescence lifetimes has also been reported in other groups[32]. Moreover, similarly decreased mean luminescence lifetime has been observed in aggregated insulin fibrils[35]. The most plausible reason for the reduced mean lifetime in aggregates could be related to changes in the fluorophore microenvironment or its interaction with surrounding charged moieties, thereby quenching the fluorophore. Moreover, the reduced mean lifetime suggests that increased steady state intensity in growing aggregates arises due to increased extinction coefficient 'ε' and not due to higher quantum yield 'Φ'. Thus, it can be stated that the change in ProCharTS luminescence lifetime both in the individual values along with their respective amplitudes can characterize SwP aggregates from their monomers.

5.2.5.2. Maximum Entropy Method (MEM) analysis

As described earlier, the time-resolved luminescence intensity decay at 340 nm was analyzed using the Maximum Entropy Method (MEM). MEM brings out the realistic contribution and distribution of the possible lifetime components along with their relative amplitudes. Here, the dataset is not confined to following any specific analysis model; instead, the decay is fitted to 100 exponential and provides the best possible fit for the decays. The luminescence intensity decays were fitted, and all the SwPs displayed trimodal distribution with 3-lifetime components, as shown in **Figure 5.10**. The randomness judged the goodness of the fit in the residuals shown in **Figure 5.11**. Similar to our Discrete analysis, the shorter lifetime components had the highest amplitude, while the longer lifetime components showed a decrease in respective amplitudes seen in **Figure 5.10**. The shorter lifetime, τ_1, is shifted more towards smaller lifetime values in aggregates than monomers. However, the intermediate and the longer lifetime components, τ_2 and τ_3, with significantly reduced amplitudes, showed nearly similar lifetime values in both the aggregates and monomers. For all the cases, the shorter component, τ_1, had luminescence lifetime values less than 1 ns, followed by the intermediate lifetime component with τ_2 in between 2-3 ns, while the longer component, τ_3, ranges between 7-10 ns. The changes are fairy conserved among all the SwPs, both in aggregates as well as in monomeric states. However, the broad distribution in the shorter lifetime component indicates the presence of multiple hidden additional lifetime subcomponents, which could not be resolved due to data acquisition limitations. Thus, the MEM analysis could distinguish the aggregates from the monomers of the SwPs and further solidify its trimodal nature and supplement the Discrete fit analysis.

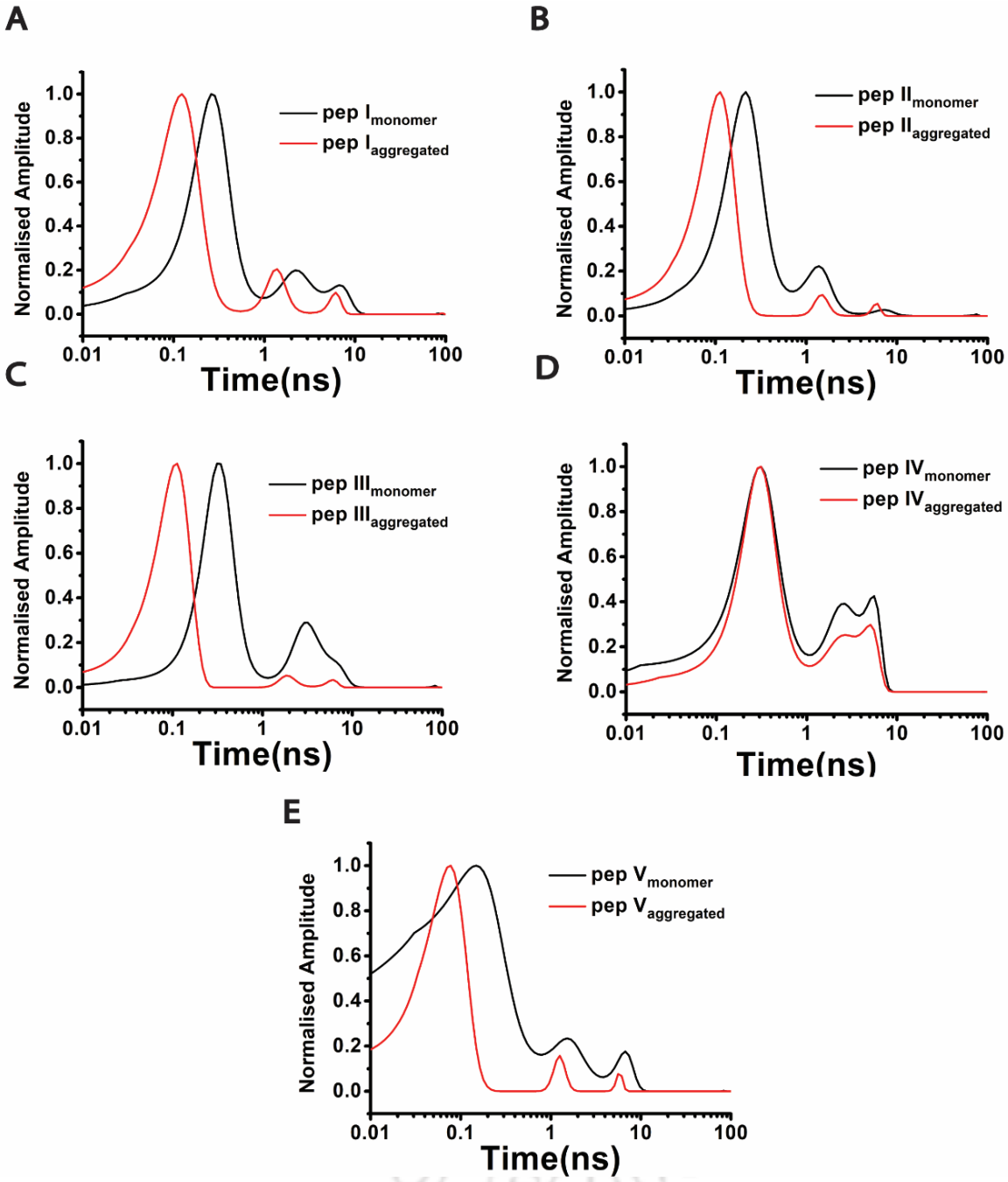


Figure 5.10: Luminescence Lifetime Distribution of SwP by MEM analysis at 340 nm excitation. Each subfigure shows the lifetime distribution of both the monomeric and aggregated states.

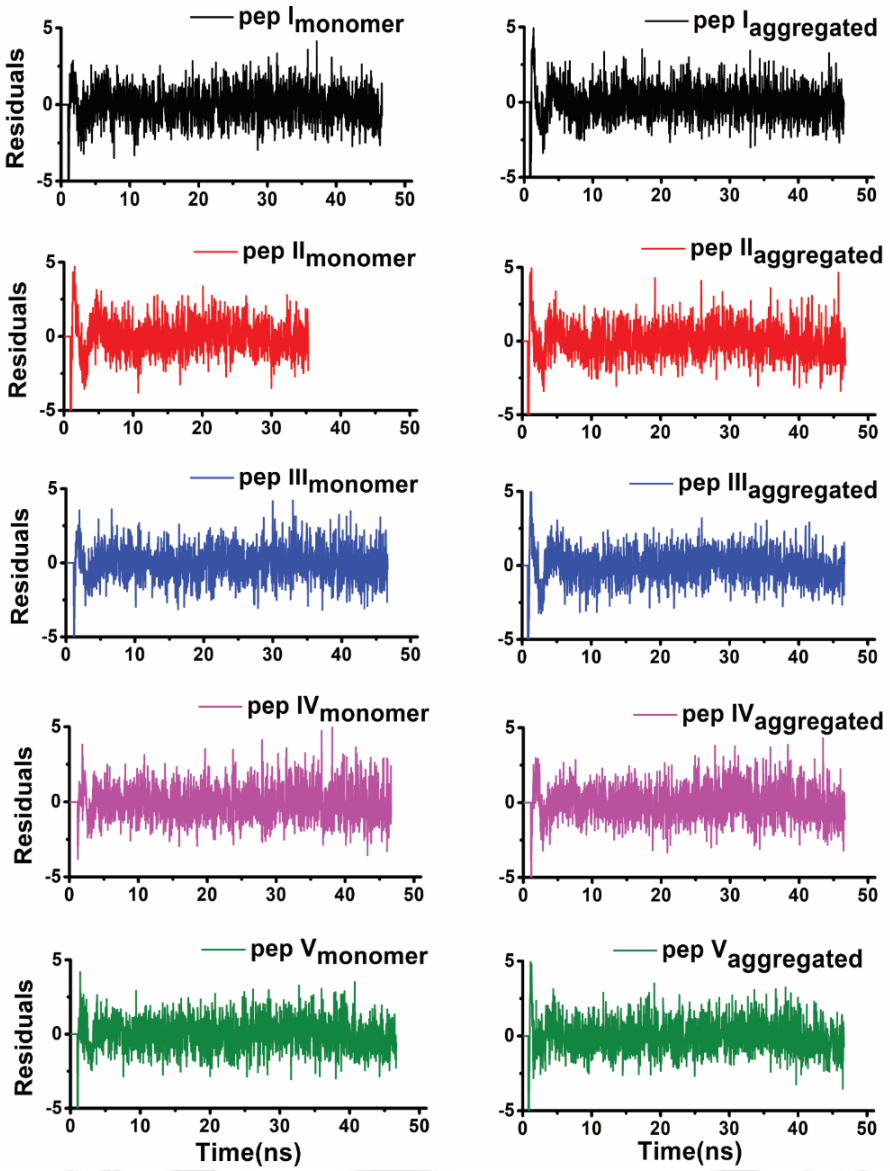


Figure 5.11: Residuals of luminescence lifetime decay for all the SwPs by MEM analysis at 340 nm excitation.

5.2.6. Comparison of the techniques

In this chapter, aggregation of SwP peptides have been monitored by multiple techniques. Next, we compared each of the techniques we have used thus far to track the early events of SwP aggregation and find out the sensitivity of each method. To compare ProCharTS absorbance and luminescence with the other conventional techniques: CD, ANS, and ThT, we calculated the percentage change in each of the methods at definite wavelengths. The difference was measured in terms of change in values of [(aggregated-monomeric)/monomeric] x 100%. The values at the start of the aggregation, i.e., 0 hr/min, defined the monomeric values, whereas for aggregated, the values at the end data point were considered in each technique. For CD, the

ellipticity values at two different wavelengths of 220 and 202 nm were used. For ProCharTS absorbance and luminescence, the respective absorption and integrated luminescence values at 340 nm wavelength were used. However, for ANS and ThT, their respective luminescence integrated areas under the curve for the aggregated and monomeric samples were considered. For Lifetime$_{340}$, the mean luminescence lifetime values were used in the analysis.

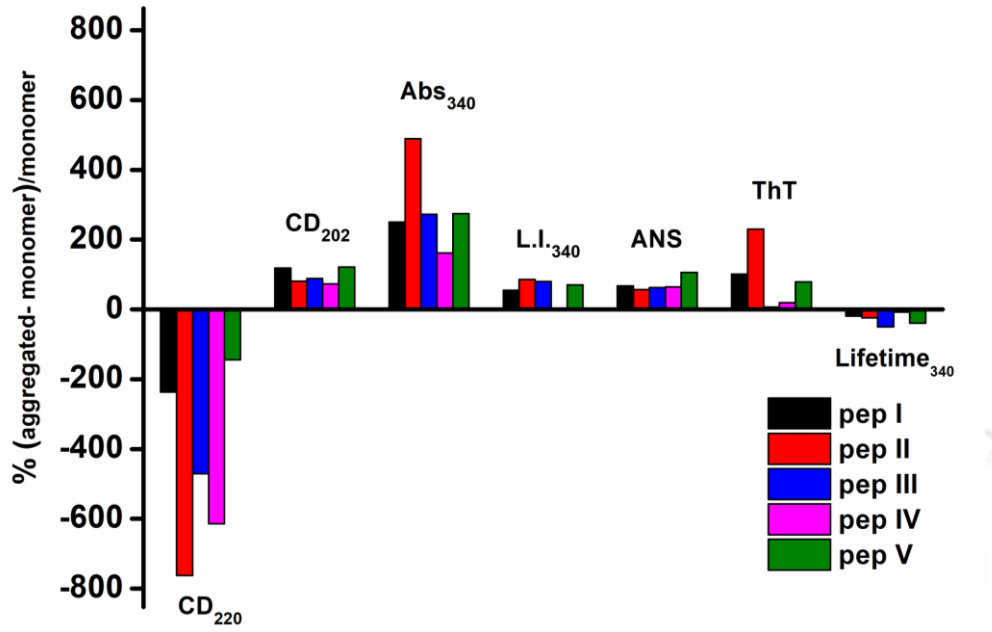


Figure 5.12: Comparison of the percentage change in the aggregated and monomeric states of SwPs for the various techniques. CD_{220}, and CD_{202}, define the percentage change of CD data at 220 nm and 202 nm, respectively, whereas Abs_{340}, $L.I._{340}$, and $Lifetime_{340}$ represent the percentage change in ProCharTS absorbance, integrated steady-state ProCharTS luminescence, and their mean luminescence lifetime on excitation at 340 nm, respectively.

As seen in **Figure 5.12**, the percentage change was found to be the most for CD_{220} and ProCharTS absorbance$_{340}$. For most SwPs, around 200% change was observed and reached as high as 500% in some peptides. However, for the other techniques, the difference was below 100% for most of them and the least observed for Lifetime$_{340}$. It is to be noted that the negative values indicate the reduced values of the aggregated samples in comparison to the monomeric states. Interestingly, the sensitivity calculated by CD_{220} and CD_{202} were different, although both are based on the same technique. The CD values at wavelength 220 nm measure the propensity of β-sheet, whereas the values at 202 nm define the random coil conformations. The incremental change at 220 nm was higher and had better sensitivity than at 202 nm. Thus, it is concluded that ProCharTS absorbance is as good as a technique as CD, if not better, to

characterize the early events of aggregation in terms of sensitivity and provides better information than ANS and ThT assays.

5.2.7. Modelling the SwP aggregation pathway

Thus far, we have analyzed each technique separately and tried to monitor the kinetics for the early events of SwP aggregation through ProCharTS and other methods. Here, we attempted to quantify the rate of the aggregation kinetics by introducing the term $t_{1/2}$. It defines the half-life of the aggregation and represents the time at which half of the aggregation is complete. We modelled the kinetics of SwP assembly and generated $t_{1/2}$ ($\nu = \frac{A_2}{2}$) in Boltzmann equation[148,149], which is shown below-

$$\nu = A_2 + \frac{A_1 - A_2}{1 + e^{(t-t_0)/dt}} \qquad (5.1)$$

where ν is the observable value, A_1, A_2, and t_0 are constants.

The aggregation kinetics of the SwPs as already observed for CD (**Figure 5.1.** ***Insets***), ProCharTS absorbance (**Figure 5.4.B, D, F, H, J**), and ProCharTS luminescence (**Figure 5.7.D, H, L, P, T**) were used. The aggregation curves fitted mostly as a sigmoid ($R^2 > 0.90$).

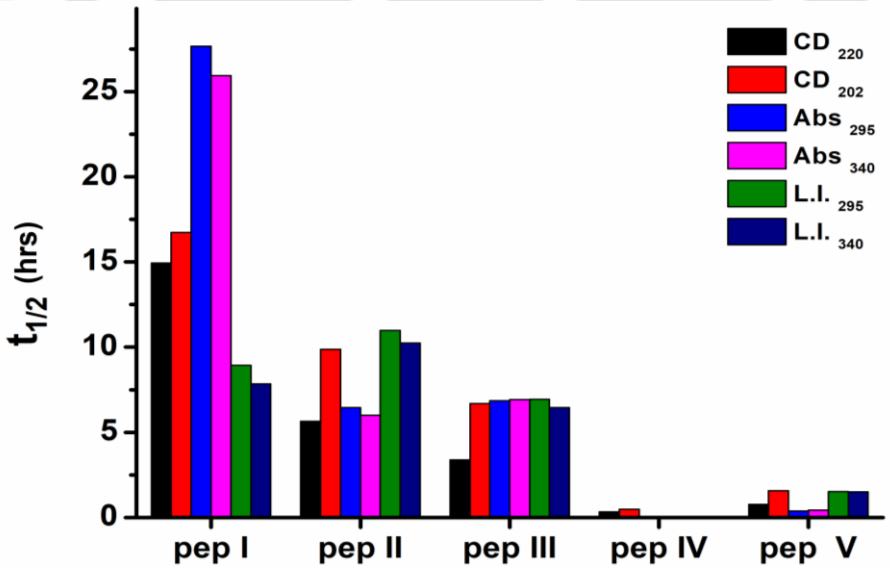


Figure 5.13: Half-life of the aggregation kinetics of SwPs obtained in CD, ProCharTS absorbance, and ProCharTS luminescence. CD_{220}, and CD_{202}, represent $t_{1/2}$ of aggregation kinetics obtained from CD data at 220 nm and 202 nm respectively. Abs_{295}, and Abs_{340}, represents $t_{1/2}$ of aggregation kinetics obtained from ProCharTS absorbance at 295 and 340 nm, and $L.I._{295}$, $L.I._{340}$, indicate the same for integrated ProCharTS luminescence, respectively.

As shown in **Figure 5.13**, the $t_{1/2}$ obtained from the CD_{220} showed the fastest aggregation in that pep IV (~0.33 hrs), followed by pep V (~0.77 hrs), pep III (3.37 hrs), pep II (~5.64 hrs), and slowest for pep I (~15 hrs). Similarly, the $t_{1/2}$ for CD_{202} was calculated, but the values were marginally higher. However, the order of the kinetics was consistent in both the wavelengths, i.e., pep IV < pep V < pep III <pep II < pep I. The change in $t_{1/2}$ values at different wavelengths from the same technique points out the difference in sensitivity at the different wavelengths. For pep III, pep IV, and pep V, the $t_{1/2}$ calculated from ProCharsTS absorbance and luminescence at 295 and 340 nm showed almost similar values. However, for pep I, the $t_{1/2}$ values from ProcharTS absorbance and luminescence showed higher and lower values, respectively, compared to that from CD. For pep II, ProCharTS absorbance more or less matched the CD, whereas ProcharTS luminescence showed marginally lower values. The variation in $t_{1/2}$ values among these techniques can be expected. Different methods are based on different principles that shed light on the information related to the different molecular phenomena occurring during the aggregation kinetics. The $t_{1/2}$ values from CD inform about the kinetics for the transition from random coil to β-sheet, whereas the ProCharTS highlights how fast the charged residues are coming close. Moreover, the $t_{1/2}$ for both ProCharTS absorbance and luminescence at 295 and 340 nm are quite comparable, suggesting the measured $t_{1/2}$ is the same irrespective of the wavelength. To sum up, in this study, ProCharTS was found to be a successful technique for monitoring the early events of SwP aggregation with high sensitivity. However, more work needs to be done to establish ProCharTS as an analytical tool to study the kinetics of any aggregated proteins or peptides.

5.3. Conclusions

- ProCharTS absorbance could detect the early events of SwP oligomerization and aggregation with a long tail extending to 800 nm. ProCharTS absorbance correlated well with CD data.

- Both ANS assay and CD can detect early events of aggregation. ANS showed drastic blue-shifted spectra with an increase in intensity as the oligomerization of peptides happens in the early stages of aggregation.

- ProCharTS luminescence generally increased with aggregation but was not consistent for all peptides. Similar results were observed for the Thioflavin T assay. The acetylated peptides showed considerably more changes in ThT as compared to non-acetylated peptides

- Luminescence lifetime studies revealed 3 exponential decay characteristics of ProCharTS. The mean luminescence lifetime decreased with aggregation; the possible reason being quenching.

- The sensitivity of CD_{220} and ProCharTS absorbance was higher than the rest of the techniques.

- The aggregation kinetics of the peptides could be modelled successfully for ProCharTS absorbance and luminescence, which followed a sigmoidal profile similar to that of CD.

Chapter 6

To monitor the Hepatitis B Virus Core Protein (HBV Cp) capsid assembly in real-time using ProCharTS and extrinsic chromophores

Chapter 6

6.1. Introduction

Hepatitis B virus (HBV) is a small enveloped virus with a diameter of 40-42 nm belonging to the hepadnavirus genus. It has a double-layered morphology and predominantly infects the hepatocytes. Hepatitis B virus Core protein (HBV Cp) is an essential protein found in the HBV virus, which mainly encapsulates the pgRNA during assembly of icosahedral capsid formed by dimers of HBV Cp. Moreover, it regulates DNA synthesis, transports partially double-stranded DNA into the nucleus, and virus maturation and release. Thus, HBV Cp is one of the promising targets for new antiviral agents to cure HBV epidemiology. To understand the capsid assembly and for the initial screening of drugs, *in vitro* or *in silico* studies are often routinely used. Here, in this study, the truncated mutant HBV Cp 149 capsid assembly is studied *in vitro* characterized by a slow nucleation step (formation of a trimer of dimers), followed by a rapid elongation phase upon addition of NaCl. Weak dimer-dimer contacts are sufficient to initial the capsid assembly[123,190]. Moreover, multiple pathways characterizes the HBV Cp capsid assembly and the assembly kinetics is highly dependent on ionic strength and temperature[191–193]. In fact, rapidly forming capsid intermediate in the early stages of assembly is found to dictate the assembly pathway[194].

We attempted to use the novel ProCharTS absorbance and luminescence to analyze the assembly kinetics in HBV Cp dimers. It is based on a simplistic, cost-effective approach and does not require careful pre-processing or high infrastructure facilities. Instead, a portable UV Visible spectrophotometer or fluorimeter is used to monitor the viral assembly process. Moreover, the information about the molecular level contacts among key charged residues can be ascertained from ProCharTS. Furthermore, unlike many fluorescent-based techniques, it is non-invasive and does not interfere with the capsid assembly process. HBV Cp capsid assembly formation can be considered a systematic aggregation mechanism where each HBV Cp dimer arranges itself in icosahedral symmetry and comes in close contact. Each HBV Cp monomer has 4 Tryptophan and 3 Cysteine residues, out of which two exist as free cysteines, while one is involved in a disulphide bond stabilizing the dimer. It is a moderately charged protein with 20% charged content with 8 Arginine, 7 Aspartate, 2 Lysine, 10 Glutamate and 4 Histidine. Thus, it presents an ideal situation to study ProCharTS, which depends on the proximity between charged amino acid residues during capsid assembly formation.

6.2. Results and Discussions

6.2.1. Purification of HBV Cp 149

HBV Cp 149 has been expressed in *E. Coli* BL21 codon$^+$ (DE3) strain, and the purification protocol has been elaborately described in *Chapter 2*. Here, the results obtained during the purification of HBV Cp 149 are highlighted. HBV Cp was initially separated from most other proteins through ammonium sulphate precipitation. The HBV Cp 149, which was expressed in the *E. Coli* strain, remained predominantly in its capsid form. Thus, the chromatograph (shown in black line) shows that most of the protein collected in the void fractions of the SEC column mainly consisted of the HBV capsids with some impurity at lower molecular weight along with

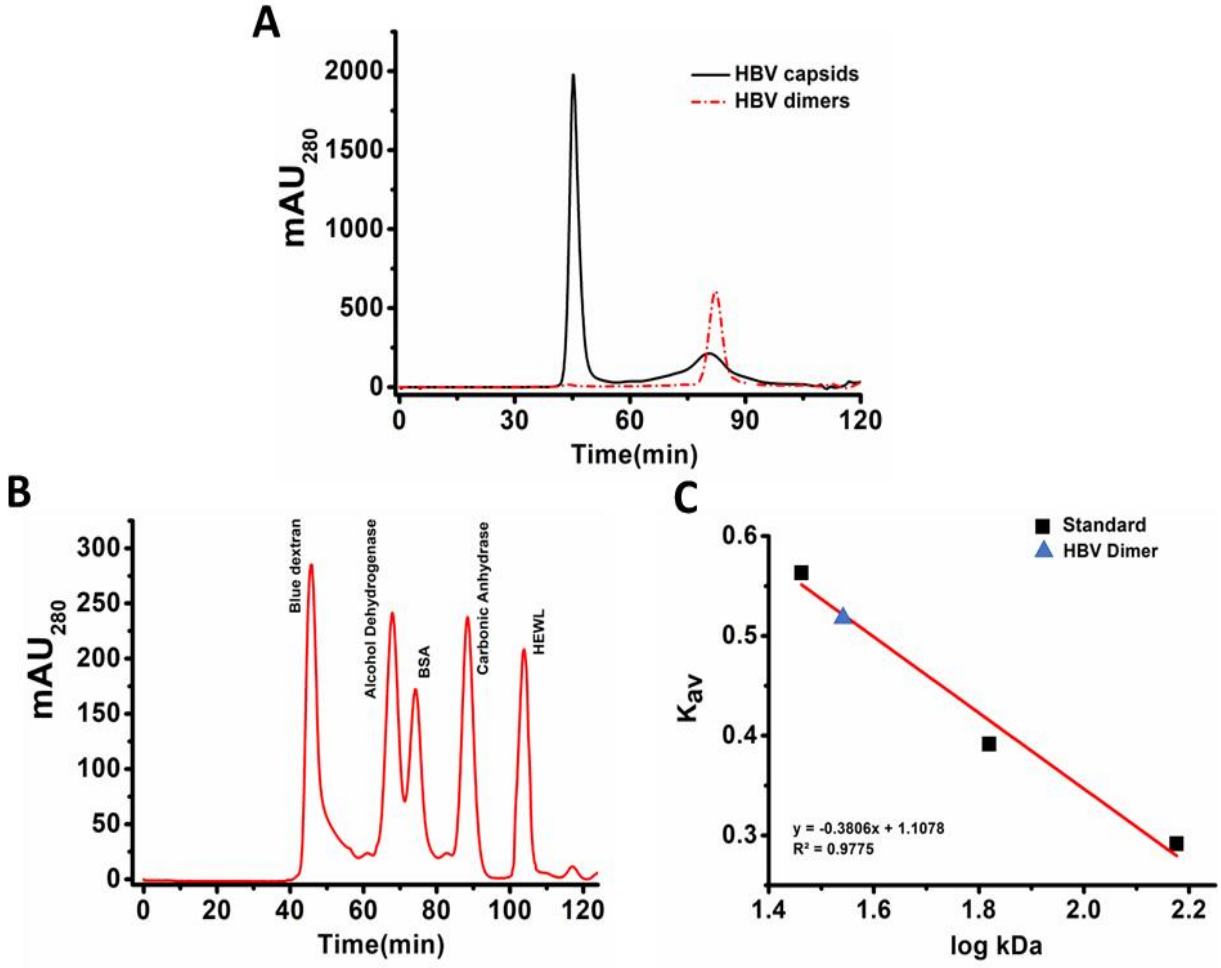


Figure 6.1: The purification of HBV Cp 149 by Size Exclusion Chromatography is shown in subfigure A. The elution volumes of various standard proteins (B) and the determination of the molecular weight of HBV Cp dimer (C) are shown, respectively. mAU_{280} represents absorbance at 280nm.

some dimers, which could be seen from 60-100 minutes of elution time as shown below in **Figure 6.1A**. Subsequently, HBV Cp 149 capsids were disassembled by 3.5 M urea and fed into the SEC column. The urea concentration was sufficient to disrupt the weak dimer-dimer contacts without disrupting the secondary structure of the dimers. The dimers eluted at longer elusion times at around 82-83 min (shown in red line), where all the impurities at higher molecular weight separated. Dimers thus collected were stored at -80°C for capsid assembly analysis.

For further characterizing the HBV Cp dimer, protein standards of different molecular weights were used, and their elution times were recorded. The standards used were: alcohol dehydrogenase (150 kDa), bovine albumin serum (66 kDa), carbonic anhydrase (29 kDa), and hen egg-white lysozyme (14.3 kDa) and the elution volumes observed for them were 67, 75, 87 and 103 mL respectively as shown in **Figure 6.1B.** Also, the column's void volume was characterized by running Blue Dextran, which is 2000 kDa and eluted at around 45 mL. The molecular weight for HBV Cp 149 dimer was derived by plotting the partition coefficient (K_{av}) vs log (kDa). The calculated molecular weight was estimated to be ~34.4 kDa, shown in **Figure 6.1C**, which roughly matches the molecular weight of the dimer.

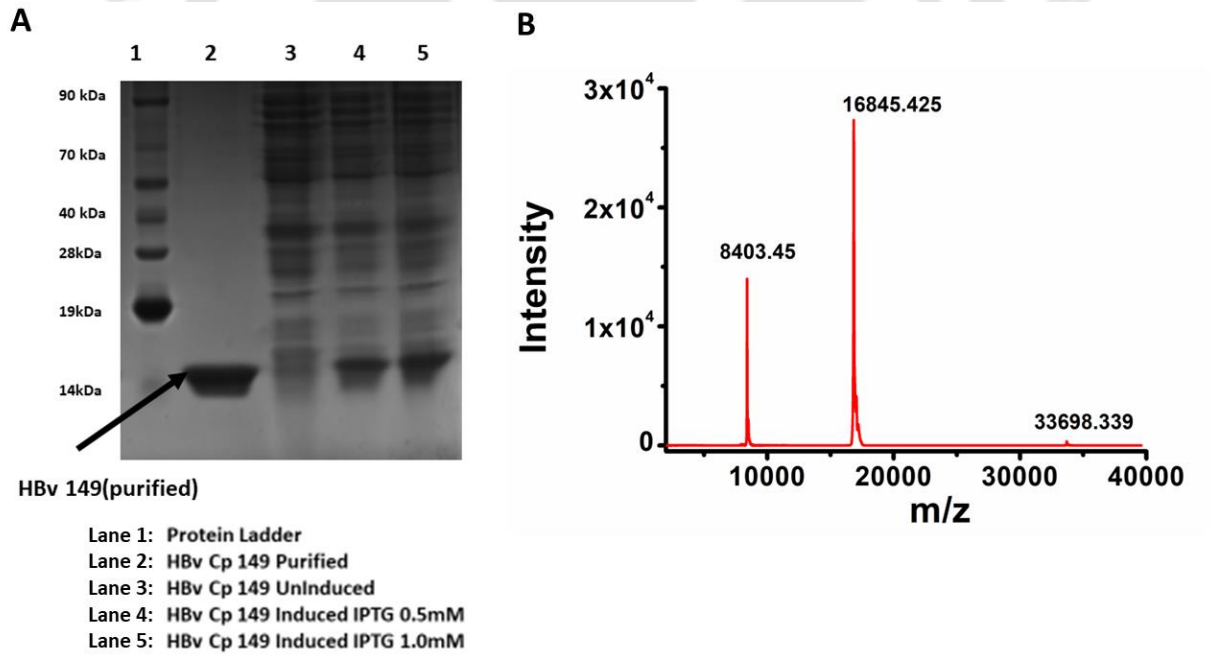


Figure 6.2: The Gel Image of expression and purification of HBV Cp 149 in 15% SDS-PAGE is shown in Subfigure A. Subfigure B shows the MALDI data for the purified protein. The expected molecular weight of HBV Cp 149 monomer is 16852.36 Da.

The HBV Cp dimers thus collected were run in 15% SDS-PAGE gel to check for purity. The gel image of the expression and purification of the protein has been shown in **Figure 6.2A**. It can be seen that in lanes 4 and 5, the HBV Cp 149 got expressed (a distinct band ~ 17 kDa) under all induced conditions. However, under the uninduced condition where no IPTG was added, a faint band at ~17 kD was visible (lane 3). The other proteins got similarly expressed in both the induced and uninduced conditions. The purified HBV Cp 149 protein could be seen in lane 2 with a clear, distinct single band. MALDI determined the exact mass of the protein to validate the purity of the HBV Cp 149. The m/z value (z = 1) for the MALDI, as shown in **Figure 6.2B,** was found to be 16845.42 Da. This calculated mass is very close to that of the theoretical value of 16852.36 Da. The peak at 8403.45 is the m/z value for z = 2 of the same.

6.2.2. Capsid assembly

The purified HBV Cp 149 dimer was dialyzed at 20 °C for 4 hrs against 50 mM HEPES 2 mM DTT, pH 7.5. The concentration was calculated to be 45 µM and was used as stock for all the

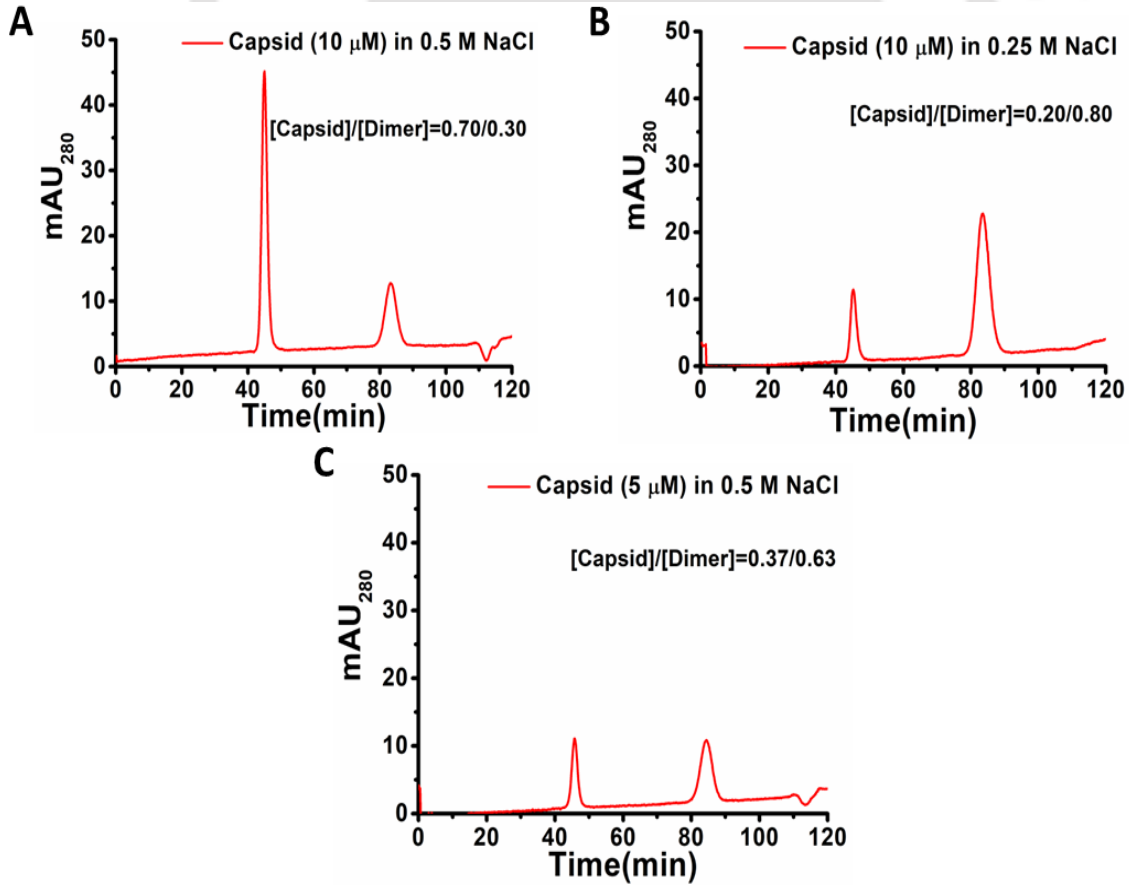


Figure 6.3: Size Exclusion Chromatographs for capsid assembly at dimer concentration of 10 µM in 0.5 M NaCl (A), 10 µM in 0.25 M NaCl (B) and 5 µM in 0.5 M NaCl (C) are shown, respectively.

experiments. Capsid assembly was studied by dissolving an appropriate amount to 10 μM in 50 mM HEPES plus 0.5 M NaCl, respectively, to initiate reaction and reaction was allowed for 30 minutes. The assembly for 5 μM in 50 mM HEPES plus 0.5 M NaCl and that for 10 μM dimer in 0.25 M NaCl were also checked. The percentage of capsids formed for 10 μM dimer in 0.25 M, 5 μM in 0.5 M NaCl and 10 μM dimer in 0.5 M NaCl were 20%, 37% and 70%, respectively, as shown in **Figure 6.3 (A, B, C).** The capsid to dimer ratio was obtained by integrating the area from the SEC chromatographs with proper baseline corrections. Thus, it could be said that NaCl accelerated the capsid assembly and increasing initial dimer concentration further increased the final capsid formation. All further study for HBV Cp assembly from here on mainly considers the capsid formation by 10 μM in 0.25 M and 0.5 M NaCl unless mentioned otherwise.

6.2.3. HBV Cp dimer and capsid structural analysis

To ensure whether the capsids formed during assembly were in the icosahedral geometry with a correct native fold, we performed Circular Dichroism. The capsids collected in the purification process in the void volume [Capsid], along with the dimers [Dimer] as well as the capsids we assembled by adding NaCl [Assembly], all showed predominant α-helix, as shown in **Figure 6.4A.** We did not see structural changes in the respective structures because the HBV assembly is a controlled accumulation of individual dimers and is not random aggregates. Moreover, the normalized ellipticity was shown in **Figure 6.4B**, which also showed that no

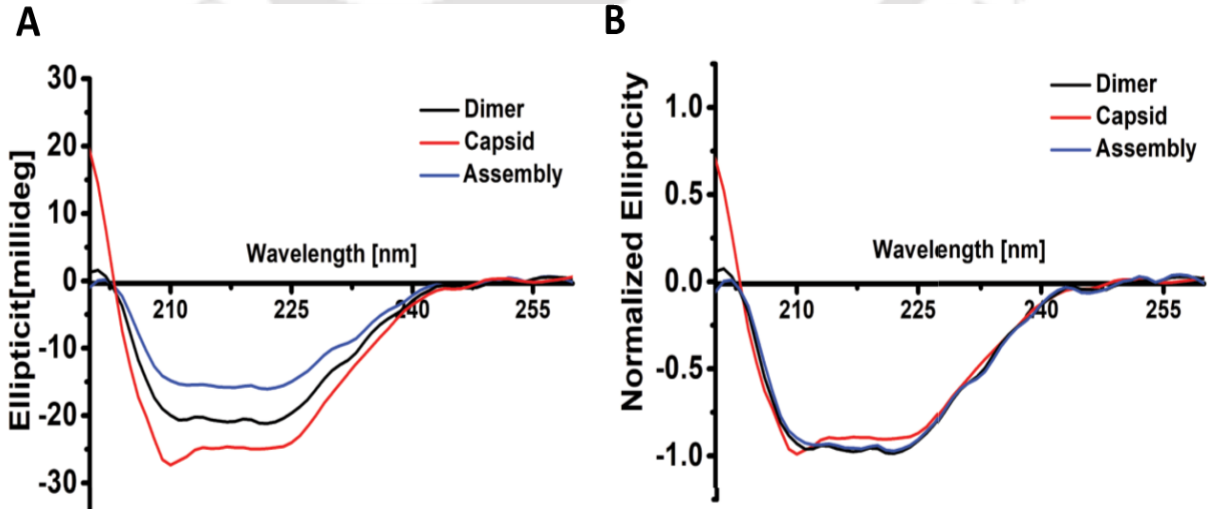


Figure 6.4: Subfigure A shows CD of the capsids (2 μM) collected in the purification process in the void volume in Tris-HCl (pH 7.4), along with the dimers (2 μM) in HEPES (pH 7.4), the capsids (1.8 μM) assembled by adding 0.5 M NaCl. Subfigure B shows their respective normalized CD spectra.

conversion from the native α helix occurred during the purification and assembly. The changes in CD intensity, as seen in **Figure 6.4A,** can be expected since they were in different buffers and with slight differences in dimer concentrations.

6.2.4. HBV capsid morphology analysis

The morphology of the HBV capsids was observed using Atomic Force Microscopy (AFM), Dynamic Light Scattering (DLS) and Transmission Electron Microscopy (TEM) for the capsids assembled in 50 mM HEPES with 0.5 M NaCl. In **Figure 6.5**, we observed the distribution of HBV capsid structures with shapes predominantly in spherical geometry.

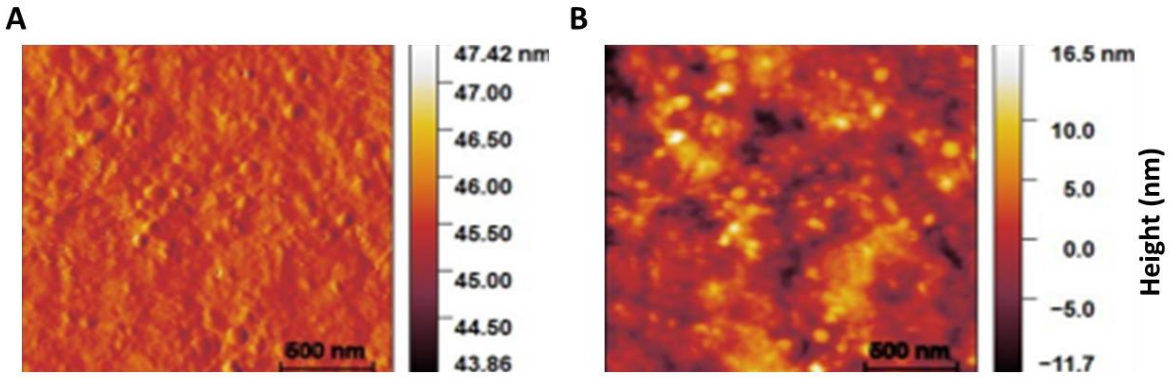


Figure 6.5: AFM of capsids (10 μM) in 50 mM HEPES with 0.5 M NaCl in non-contact mode.

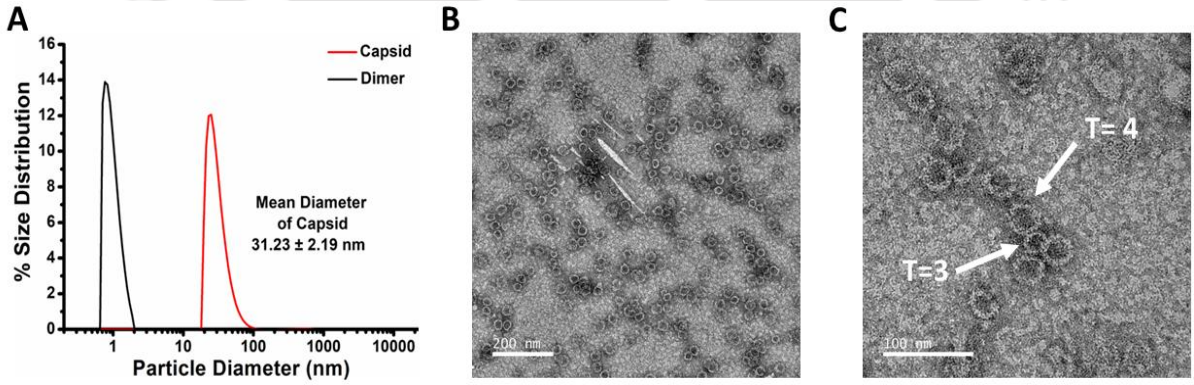


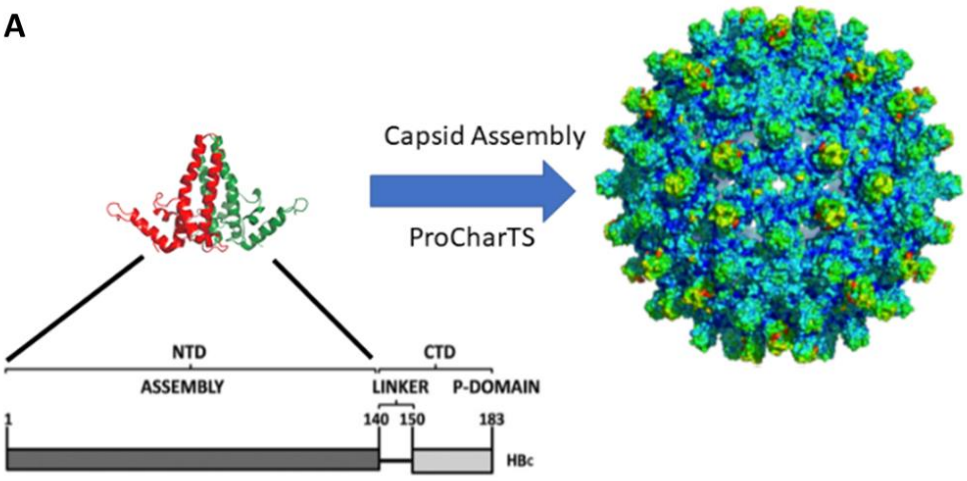
Figure 6.6: The Size Distribution by DLS (A) and morphology of HBV capsid by TEM (B, C) are shown, respectively.

Although AFM could not provide any conclusive evidence of the distribution of capsid size distribution and morphology, we subsequently performed DLS and TEM. In DLS, the distribution of the capsids was evident, with a mean diameter of 31.23 ± 2.19 nm (**Figure 6.6A**) which roughly matches the mean diameter of T=3 and T=4 capsids. The HBV dimer showed a

distribution around 2 nm, far less than that of capsids. In TEM, both T=3 and T=4 capsids with a diameter of around 30-34 nm with perforated pores distributed throughout the capsid can be observed, as seen in **Figures 6.6B and C.**

6.2.5. ProCharTS absorbance

Thus far, we have only characterized the capsid assembly under various conditions and confirmed the capsid formation. Next, we used our novel ProCharTS to investigate the capsid assembly process. During HBV Cp capsid assembly, 120 dimers arrange in an icosahedral geometry (T=4) with a diameter of around 34 nm, shown in **Figure 6.7A**. From the PDB structure of HBV Cp149 (1QGT), we constructed a distance matrix of the charged amino acid residues. It was found that 44 such contacts in each dimer of dimers are possible among charged atoms within 10 Å in monomer, dimer interface and dimer-dimer interface, which is shown in **Figure 6.7B**. Such additional contacts during assembly formation make the study of HBV capsids an ideal system for ProCharTS, which is dependent on the 3D proximity of the charged amino acids.



B

Residue Contacts (in Å)	Within a Cp 149	Intra Cp 149 dimer	Inter Cp 149 dimer
D< 5 Å	3	7	2
5 Å < D< 10 Å	15	12	5

Figure 6.7: Subfigure A shows the HBV Cp dimer and capsid structure. Subfigure B shows the calculated distance of closely spaced charged residues in monomer, dimer interface and dimer-dimer interface within 10 Å.

Here, the capsid assembly from 10 μM Cp dimers in 50 mM HEPES with 0.25 M NaCl and 0.5 M NaCl were analyzed. Capsid assembly under similar conditions were previously reported to be over within 30mins[124]. The ProCharTS absorbance for the entire capsid assembly was thus recorded at 2 min intervals, and the values for the initial (t = 0 min) and final (t = 30 mins) absorbance were shown in **Figure 6.8.** It can be seen that during capsid assembly, enhanced ProCharTS absorbance could be observed, and the changes were prominent at longer wavelengths, i.e., after 315 nm. Although slight differences in ProCharTS could be observed below 315 nm, as shown in the Inset, the Tryptophan absorbance completely masked the ProCharTS. Reduced change in ProCharTS may be due to the addition of NaCl which diminishes ProCharTS.

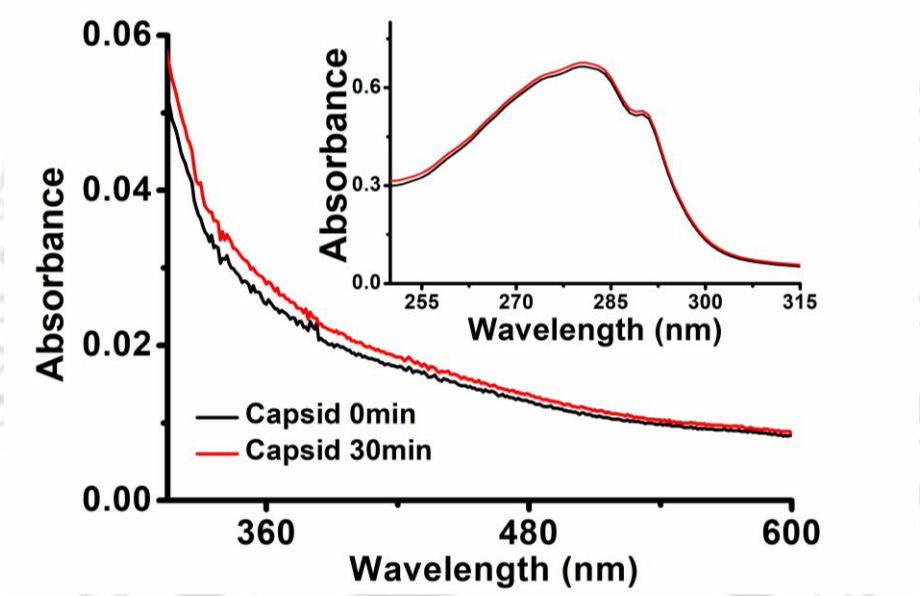


Figure 6.8: ProCharTS absorbance of values for the initial (t = 0 min) and final (t = 30 mins) observed during capsid assembly. Inset shows the change in Tryptophan absorbance below 315 nm for the same.

The changes in the ProCharTS absorbance observed in capsids for the various wavelengths: 310, 325, 340, 355, 370 and 400 nm for both dimers and capsids are shown in **Figure 6.9.** As the capsids assembled, a steady increase in ProCharTS could be observed, and the change was consistent for all the wavelengths. Dimers showed least difference in ProCharTS absorbance, whereas the capsids in both 0.25 M and 0.5 M NaCl showed a profound increase in absorption. The changes in 0.5 M NaCl were significantly greater than in 0.25 M NaCl. We further plotted the changes in ProCharTS at each wavelength: 325, 340, and 355 nm to compare the dimers and capsids, which too showed similar results as observed in **Figure 6.10.** ProCharTS absorbance decreases drastically as wavelength increases. The increase in ProCharTS signal

for every time increment (2min) fitted to a perfect sigmoidal curve below 355nm. However, at longer wavelengths greater than 400nm, the change in ProCharTS decreases to nearly one-tenth of the magnitude as compared to 295nm. Thus, the sensitivity or signal to noise ratio (S/N) of ProCharTS to monitor the sigmoidal curve falls off with increasing wavelength and were subsequently discarded in this study. Thus, it can be said that ProCharTS absorbance could detect both the very fast kinetics of assembly formation observed in 50 mM HEPES 0.5 M NaCl and the comparatively slower kinetics in 0.25 M NaCl. ProCharTS absorbance got saturated within 30 minutes in the case of 0.5 M NaCl, whereas for that of 0.25 M NaCl, the assembly was not entirely over.

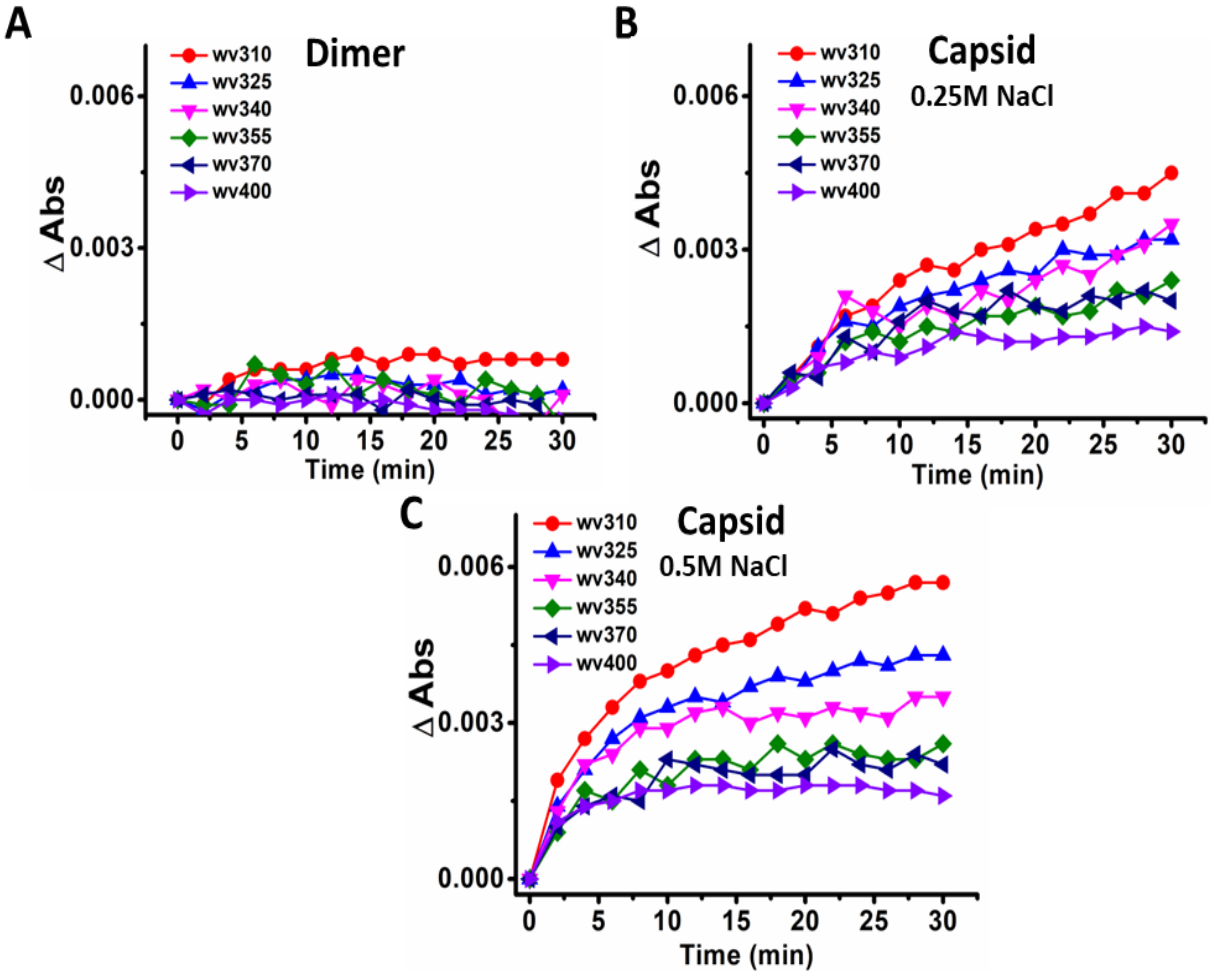


Figure 6.9: Kinetics of HBV capsid assembly by ProCharTS absorbance in 10 μM dimers (A), capsid formed the dimers in 50 mM HEPES with 0.25 M NaCl (B) and 0.5 M NaCl (C), respectively. Various wavelengths used were: 310, 325, 340, 355, 370 and 400 nm.

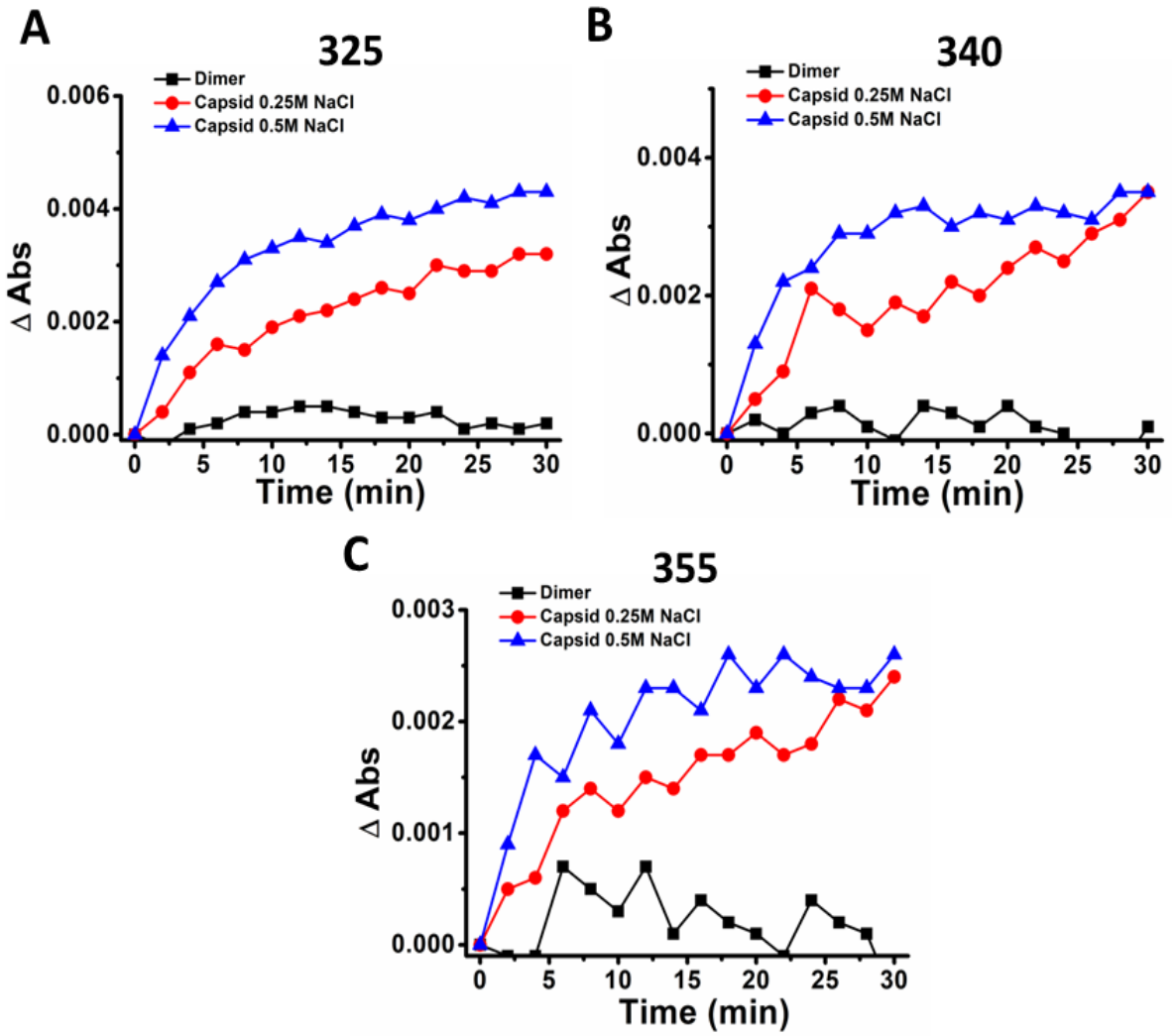


Figure 6.10: Kinetics of HBV capsid assembly by ProCharTS absorbance for different wavelengths 325 nm (A), 340 nm (B) and 355 nm (C), respectively.

To better understand the kinetics rate in the capsid formation, we tried to fit the ProCharTS absorbance using **Equation 6.1**. Since the capsids formed in 50 mM HEPES, 0.25 M NaCl was sufficiently less than 0.5 M NaCl **(Figure 6.3B).** We only used the ProCharTS absorbance in 50 mM HEPES 0.5 M NaCl for the fitting and all the subsequent studies.

Capsid Assembly kinetics is represented by-

$$r = A\left(\frac{kt}{1+kt}\right)^{n} \qquad (6.1)$$

where *A* is a constant, *k* is the time constant, and *n* is the exponential term.

No simple model can adequately describe the virus assembly reaction where multiple small subunits sequentially add to the growth in the capsid. However, this simplified hyperbolic curve with the exponential term 'n' approximates the kinetics of capsid with a sigmoidal fit characterized by either a long lag phase or fast initial kinetics which slowly, asymptotically approach equilibrium. This term 'n' is analogous to the Hill coefficient[141], which is often used for cooperative binding isotherm. For HBV aggregation, where the lag phase is not so prominent, it is safe to assume a value of n=1[195]. However, the initial value of n was both allowed to move freely as well as keeping the value of n fixed to 1 and comparing the fitting. Reliable kinetics were generated for n=1 under all the circumstances, as shown in **Table 6.1**. The inverse of the time constant 'k' signifies the time at which 50% of the capsid formation is over, which is referred to as $t_{1/2}$.

		A	n	k (s^{-1})	$t_{1/2}$ or 1/k (s)	R^2
n as variable	**Abs_{325}**	0.00507	0.87643	0.00253	395.25	0.99349
	Abs_{340}	0.00359	-718.825	355.6006	----	0.97651
	Abs_{355}	0.00274	1.50366	0.00816	122.55	0.92729
n = 1	**Abs_{325}**	0.00499	1	0.00315	317.46	0.99382
	Abs_{340}	0.00374	1	0.00554	180.51	0.97269
	Abs_{355}	0.00279	1	0.00463	215.98	0.93194

Table 6.1: Parameters obtained from the fitted curve for ProCharTS for n=1 and n as a variable. Abs_{295}, Abs_{325} and Abs_{355} represent the different ProCharTS absorbance at 325, 340 and 355 nm, respectively.

As seen in **Table 6.1,** the parameters for the sigmoidal curve represented better fits for all the wavelengths when n was set to 1. The fitted curves yielded good R^2 with $t_{1/2}$ values around 3-5 minutes. Such fast kinetics of capsid assembly is often observed for various other viruses[196]. The fitted curve has been shown in **Figures 6.11 A**, **B, and C,** respectively.

Thus, it can be summarised that ProCharTS absorbance can monitor the fast kinetics of HBV capsid assembly. During assembly formation, the addition of the individual dimers to the capsid forms clusters of charged residues with increased contacts, thereby enhancing the ProCharTS absorbance.

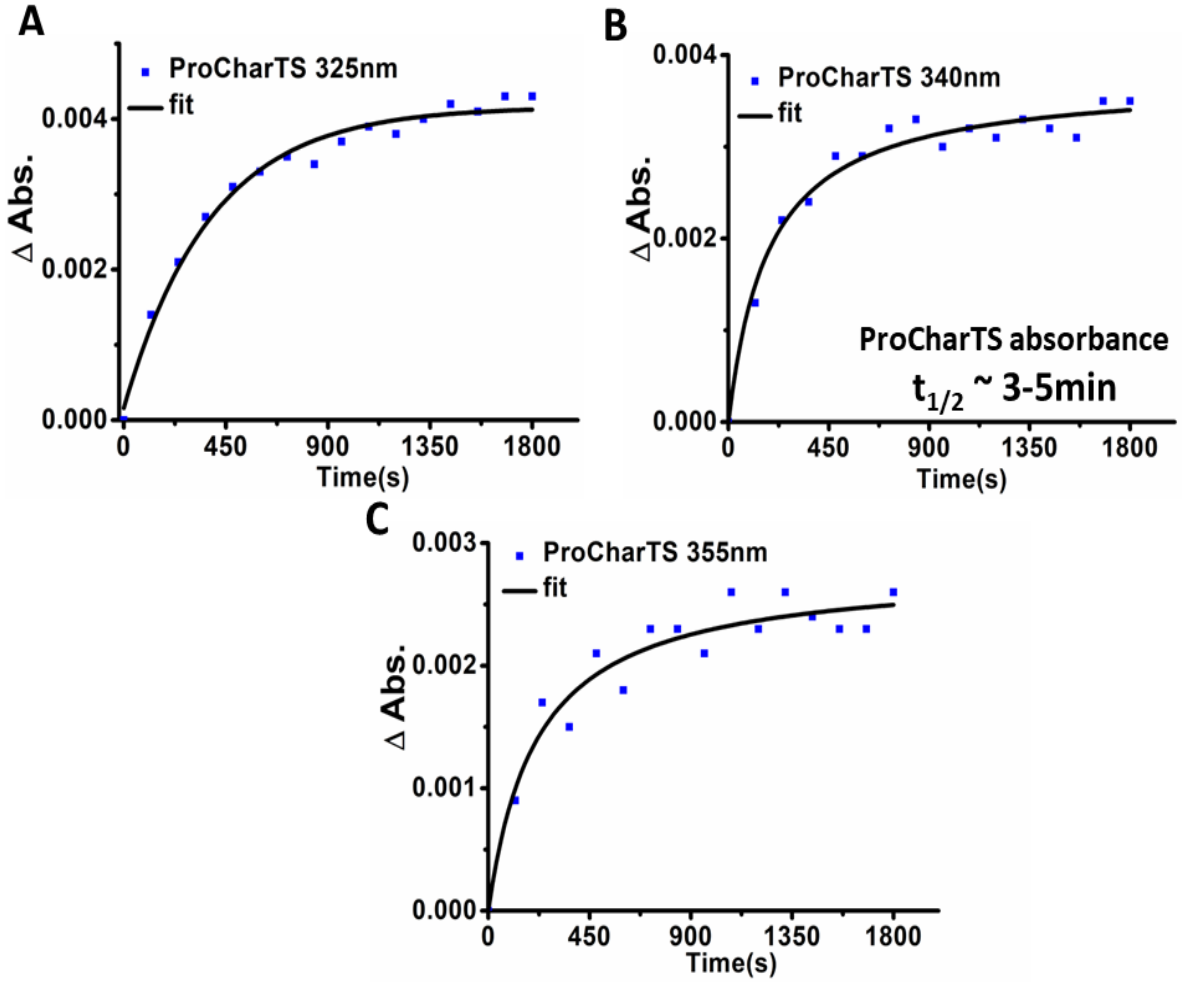


Figure 6.11: Fitted HBV capsid assembly curves for ProCharTS absorbance at 325 nm (A), 340 nm (B), and 355 nm (C) respectively. The black line shows the fit for the curve.

6.2.6. Tryptophan fluorescence and ProCharTS luminescence

Similar to ProCharTS absorbance, the steady-state luminescence properties arising from the HBV capsids were explored. Initially, the fluorescence of Tryptophan was monitored at 295 nm excitation, and the emission was collected from 310-450 nm. As the capsids were formed, the fluorescence emission maxima (λ_{max}) shifted from 323 nm (dimers) to 320 nm. The blue shift in the λ_{max} was accompanied by increased fluorescence intensity, as shown in **Figure 6.12A**. This decrease in emission wavelength and increased fluorescence intensity was primarily due to the Tryptophan's more buried position in capsids than dimers. To eliminate the possibility of increasing NaCl to the observed emission during capsid assembly, we performed a control experiment keeping NATA (a tryptophan analogue) under the same conditions. However, no change in fluorescent intensity or λ_{max} was observed, as shown in Inset. We further observed ProCharTS luminescence at various excitation wavelengths: 340,

355, and 370 nm and collected the corresponding luminescence from 370-560 nm, 380-570 nm, and 400-570 nm. Interestingly, the luminescence was more in the case of the capsids as compared to dimers, as shown in **Figures 6.12B, C,** and **D**. The area under the emission curve for both the Tryptophan fluorescence and ProCharTS luminescence was integrated and plotted versus wavelength. It was found, as seen in **Figure 6.12E,** that Tryptophan fluorescence increased appreciably, whereas the ProCharTS luminescence intensity approximately doubled in capsids. However, it should be noted that the ProCharTS luminescence intensity was nowhere close (~100 times less) to the fluorescence intensity of Tryptophan.

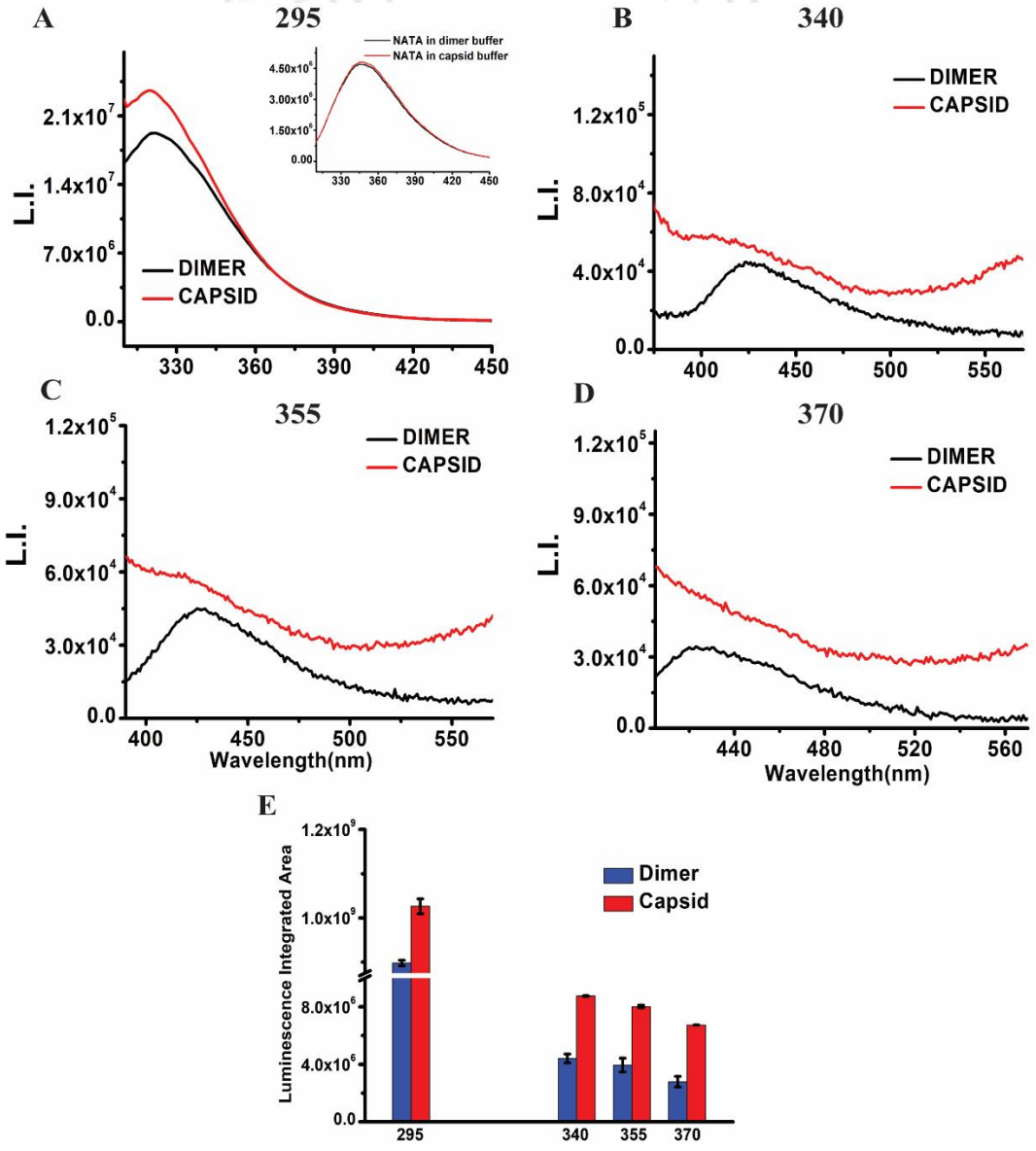


Figure 6.12: The Tryptophan fluorescence emission at 295 nm (A) and ProCharTS luminescence at 340 nm (B), 355 nm (C) and 370 nm (D) are shown. The Inset in Subfigure A shows the fluorescence from NATA. Subfigure E compares the integrated Trp fluorescence with ProCharTS luminescence at different excitation wavelengths.

Since the time required for data acquisition in scanning the entire emission spectra at an excitation wavelength in steady-state luminescence was significant compared to the assembly kinetics. The change in ProCharTS luminescence intensity was carried out at the selected excitation wavelength keeping the emission fixed at the λ_{max}. The assembly kinetics using Tryptophan fluorescence at 295 nm and ProCharTS luminescence at 340 nm excitation was subsequently carried out with emission fixed at 320 nm and 420 nm, respectively. The time interval between each acquisition was set at 0.1s, and the kinetics was analyzed for 30 minutes. We found that the increase in ProCharTS luminescence was similar to that observed for ProCharTS absorbance and followed a sigmoidal curve shown in **Figure 6.13**. Surprisingly, the enhanced ProCharTS luminescence was consistent with Tryptophan fluorescence**.** Both the luminescence was fitted to the **Equation 6.1** already discussed, and calculated $t_{1/2}$ values were around 3-5 minutes. The goodness of the fit was assessed by the decent R^2 values > 0.96. The parameters obtained from the fit are shown in **Table 6.2**. Similar $t_{1/2}$ values have also been obtained from ProCharTS absorbance, although they were not precisely the same. Although Tryptophan fluorescence has been utilized to monitor other viral capsid assemblies[197], this is the first time that ProCharTS luminescence has been used to observe and distinguish the capsids as compared to dimers.

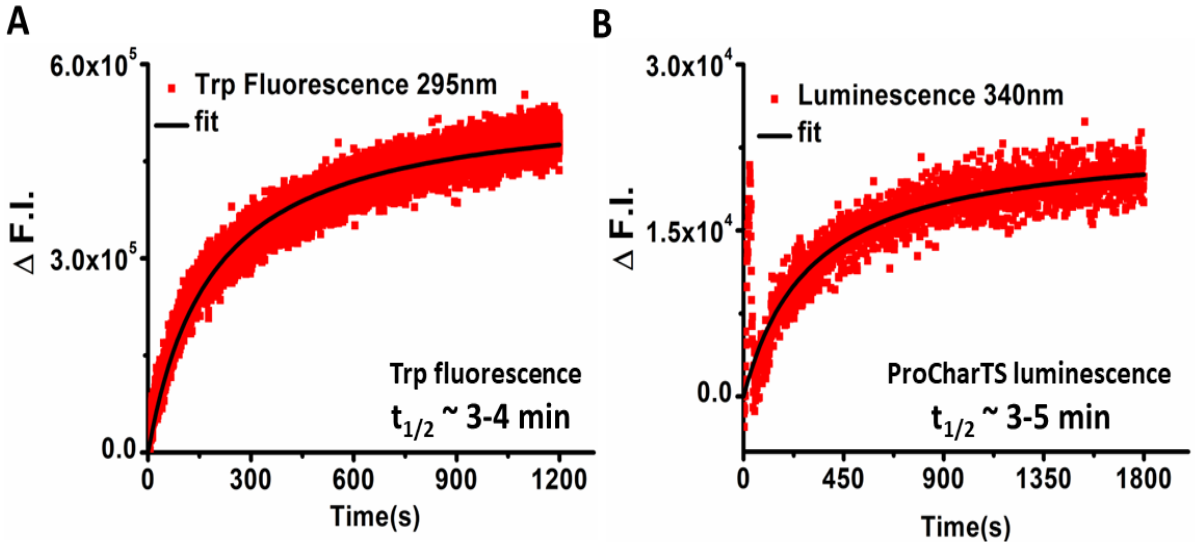


Figure 6.13: Subfigure A shows the time dependence of Tryptophan Fluorescence when excited at 295 nm and emission collected at 320 nm. Subfigure B shows the change in ProCharTS intrinsic luminescence at 340 nm excitation and emission collected at 420 nm. The black line indicates the fit for the curve.

		A	k (s^{-1})	$t_{1/2}$ or 1/k (s)	R^2
n =1	**F.I.$_{295}$**	5.492 x 10^4	0.00538	185.87	0.9648
	F.I.$_{340}$	2.340 x 10^3	0.00332	301.20	0.9830

Table 6.2: Parameters obtained from the fitted curve for ProCharTS luminescence at 295 and 340 nm for n=1. F.I.$_{295}$ and F.I.$_{340}$ represent the ProCharTS luminescence at 295 and 340 nm, respectively.

6.2.7. 90° Light scattering

90° Light scattering is often used as a gold standard to detect the formation of aggregates or any assembly or cluster formation. Here, we used scattering to assess the capsid assembly kinetics generated from ProCharTS, both qualitatively and quantitatively. Similarly, the capsid assembly was carried out at 320 nm and 340 nm separately, and the scatter was collected at 90°. The scatter for HBV Cp 149 assembly is often carried out at 320 nm, but here we also included the 340 nm excitation to compare with the ProCharTS phenomenon.

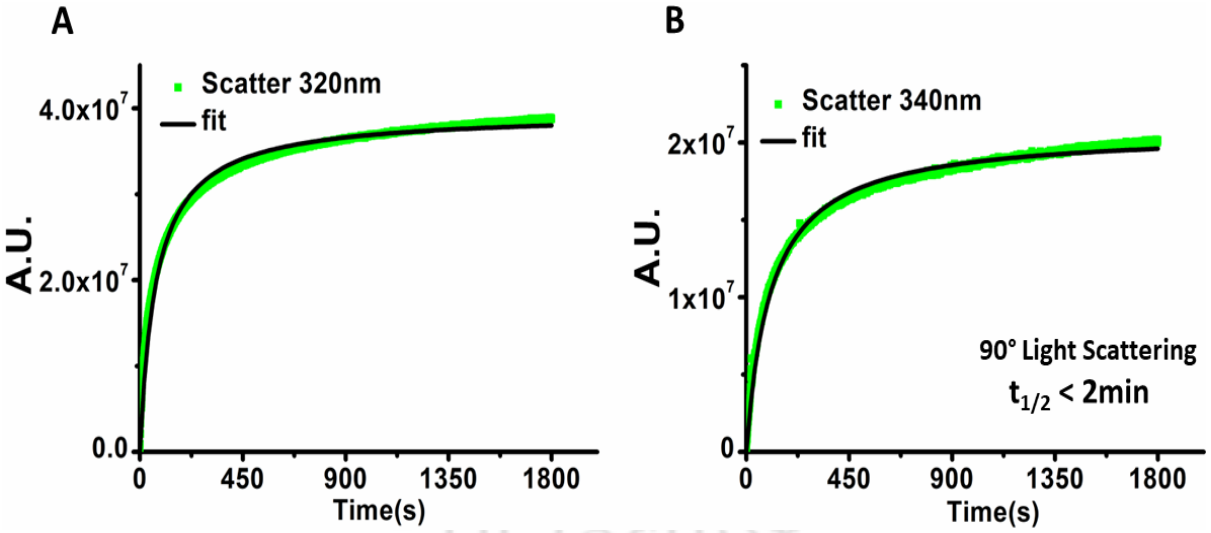


Figure 6.14: 90° light scattering curve for capsid assembly at 320 nm (A) and 340 nm (B) in real-time for capsids (10 μM) in 50 mM HEPES with 0.5 M NaCl. The black line shows the fit for the curve.

		A	k (s^{-1})	$t_{1/2}$ or 1/k (s)	R^2
n =1	**S_{320}**	3.948 x 10^7	0.01416	70.72	0.9844
	S_{340}	2.077 x 10^7	0.00929	107.64	0.9860

Table 6.3: Parameters obtained from the fitted curve for 90° Light scattering at 320 and 340 nm for n=1. S_{320} and S_{340} represent the 90° Light scattering at 295 and 340 nm, respectively.

Strikingly, the scatter too followed a sigmoidal curve that fitted to **Equation 6.1.** The R^2 values for the fit were > 0.98, and the parameters generated from the fit have been shown in **Table 6.3**. The calculated $t_{1/2}$ values were less than 2 minutes. The $t_{1/2}$ values thus showed faster kinetics as compared to both ProCharTS absorbance and fluorescence. Most probably, the more rapid kinetics observed in light scattering measured the elongation of the capsids due to the addition of each dimer subunit. However, the ProCharTS and intrinsic Tryptophan fluorescence were able to pick up the relatively later stages of maturation and subsequent reorientation of the dimers in capsids ($t_{1/2}$ ~3-5 minutes). Such reorientation of misaligned dimers and proper capsid formation in the late stages has been a characteristic feature in HBV capsid assembly [127,128]. Interestingly, similar delayed kinetics from Tryptophan fluorescence has been observed to detect the post-assembly structure modification in capsid assembly of human papillomavirus-like particles (VLPs)[197]. Thus, ProCharTS and 90° light scattering, although measuring different morphological features in HBV capsid assembly, are simple, inexpensive, label-free techniques than can accurately monitor HBV capsid assembly formation. Moreover, the ease of use and reproducibility observed from ProcharTS in characterizing such ultra-fast kinetics could eliminate the usage of extrinsic probes and sophisticated high-end technologies for rapid sensing and screening of antivirals (CpAMs) for HBV infection.

6.2.8. Labelling of HBV dimer with 1,5-IAEDANS and Dansyl Chloride

Thus far, ProCharTS and light scattering have been used to characterize HBV capsid assembly. Next, the capsid assembly is interrogated by using extrinsic fluorescent probes. The HBV dimers were labelled with: 1,5-IAEDANS and Dansyl Chloride separately. 1,5-IAEDANS was used to label the free Cysteine, whereas Dansyl Chloride was used for lysine. Each monomer of HBV has 2 Lysine and 3 Cysteines, out of which one of the cysteines are involved in di-sulphide bond formation in the HBV dimer. Both the probes were successfully labelled to the HBV dimer separately with a Dye/Protein (D/P) ratio of 0.7 and 1.4 for 1,5-IAEDANS and Dansyl Chloride, respectively. It is to be noted that the protein referred to here in the D/P ratio is not the HBV monomer but the dimer instead. For the capsid assembly formation, varying dilutions of dimers (labelled/unlabeled) were used with a final dimer concentration of 10 μM. For 1-5-IAEDANS, the labelled/unlabelled dimer ratios were 1:10 and 1:15, whereas, for Dansyl Chloride, the proportions were 1:15 and 1:30. The reason for the two-fold dilution of labelled dimer in Dansyl Chloride was due to its doubled D/P ratio. Although the labelling of

both the probes to HBV dimer was confirmed by UV-Visible spectroscopy, the labelling for 1,5-IAEDANS was further cross-checked by Föster Resonance Energy Transfer (FRET). The Föster critical distance, R_0 value is 22 Å for Tryptophan (donor) and 1,5-IAEDANS (acceptor) pair[198]. When Tryptophan was excited at 295 nm, the proximity of the labelled donor-acceptor pair is reflected in the FRET, where an enhanced acceptor fluorescence at 480 nm was observed with a respective decline in the donor fluorescence at 323 nm, as shown in **Figure 6.15.**

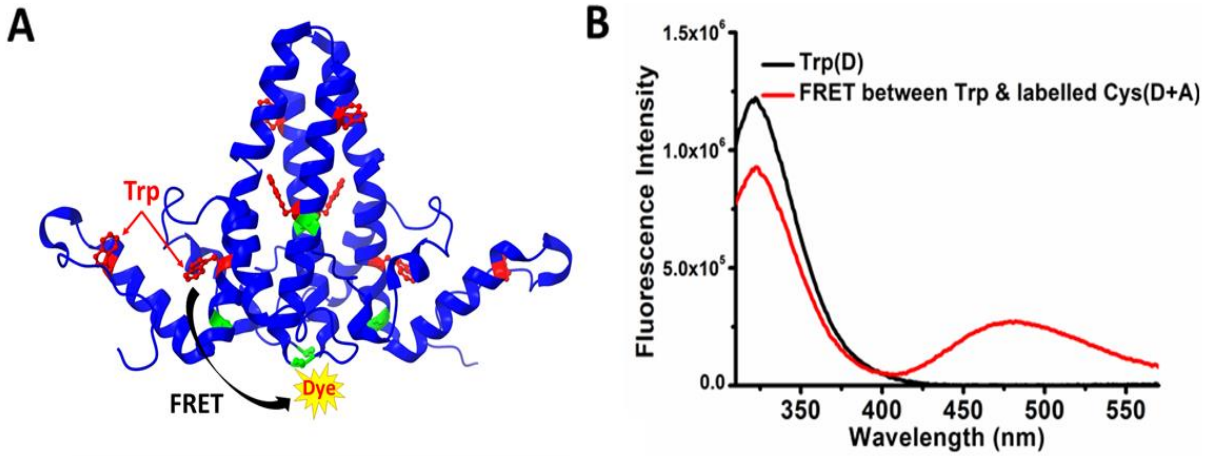


Figure 6.15: FRET observed in HBV dimer (PDB ID:3J2V) labelled with 1-5-IEADANS (Dye/Dimer:0.7) is shown in subfigure A. Subfigure B shows the changes in the fluorescence spectrum of Tryptophan in successfully tagged Cysteine during FRET in HBV Cp149 dimer.

6.2.8.1. Steady-state fluorescence of labelled HBV

The steady-state fluorescence for each case of the labelled HBV dimers was recorded. **Figures 6.16** and **6.17** show the fluorescence spectra for capsid and dimers using both 1,5-IAEDANS and Dansyl Chloride, respectively. Unlike Tryptophan fluorescence which showed a marginal decrease in fluorescence emission maxima (λ_{max}) from 323 nm to 320 nm during capsid formation, it could be seen that for both the probes and each dilution, the fluorescence emission maxima (λ_{max}) did not change at all. The λ_{max} observed for 1,5-IAEDANS, and Dansyl Chloride at 340 nm excitation was 488 ± 1 nm and 439 ± 1 nm, respectively.

However, minor changes in the fluorescent intensity could be observed. This small change could be due to the subtle differences in the local environment around the probes during capsid formation. Next, we proceeded to characterize the capsids using steady-state fluorescence anisotropy.

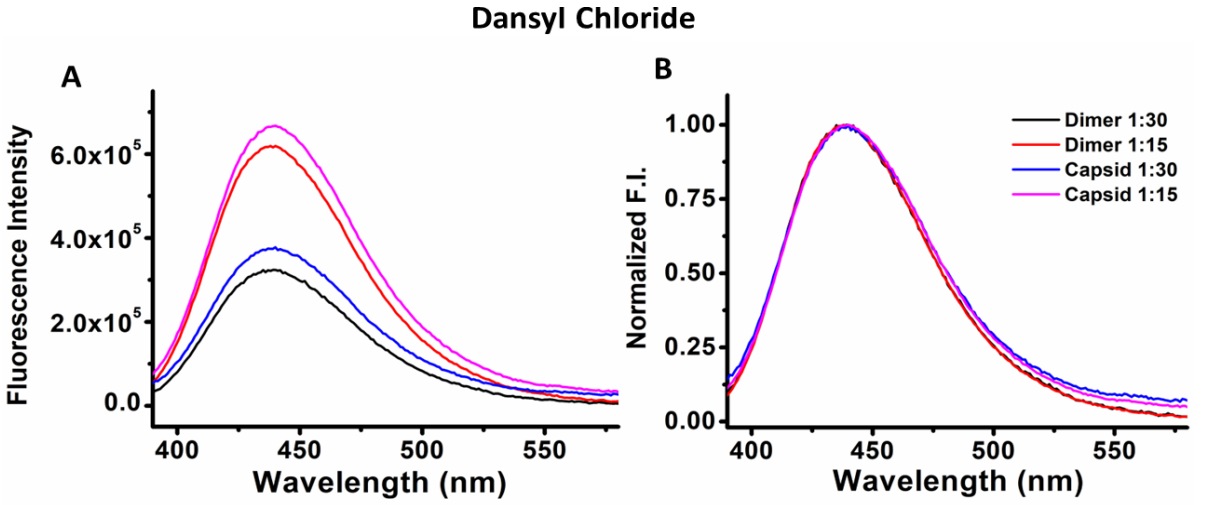


Figure 6.16: Steady-state Luminescence intensity (A) and Normalized Intensity (B) of Dansyl Chloride labelled Dimers and Capsid of HBV

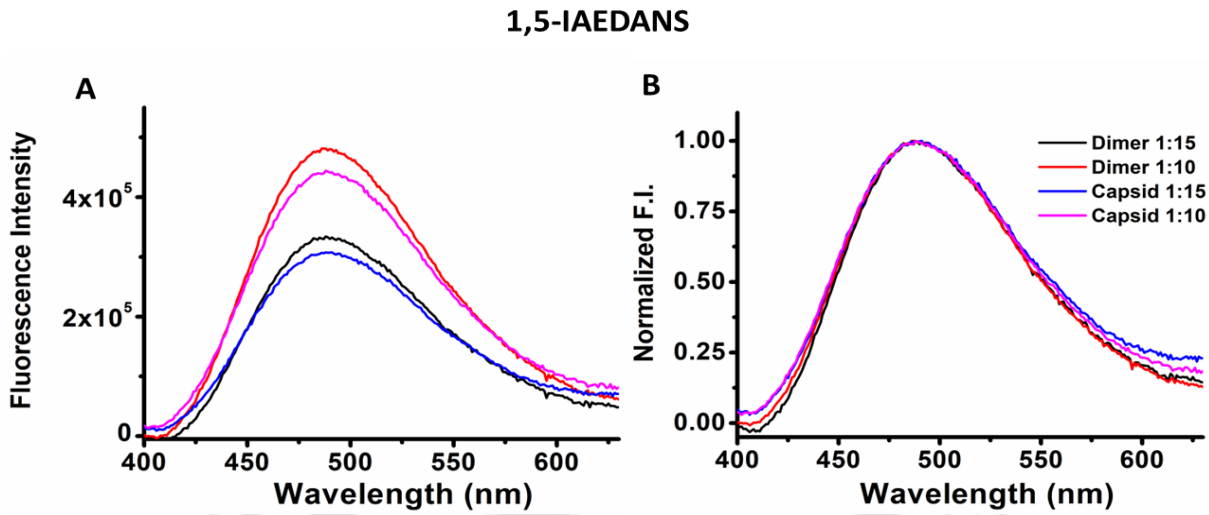


Figure 6.17: Steady-state Luminescence intensity (A) and Normalized Intensity (B) of 1,5-IAEDANS labelled Dimers and Capsid of HBV

6.2.8.2. Steady-state fluorescence anisotropy of Tryptophan and labelled HBV

The steady-state fluorescence anisotropy for both the probes: 1,5-IAEDANS and Dansyl Chloride, was carried out along with that for Tryptophan. Steady-state fluorescence measures the magnitude of depolarization through rotational motion. It thus gives essential information on the rotational diffusion of the macromolecule to which the fluorophore is attached. The steady-state anisotropy can provide insights into the macromolecule's size and shape as long as the fluorescence lifetime of the fluorophore is in the range of its rotational tumbling time of the macromolecule.

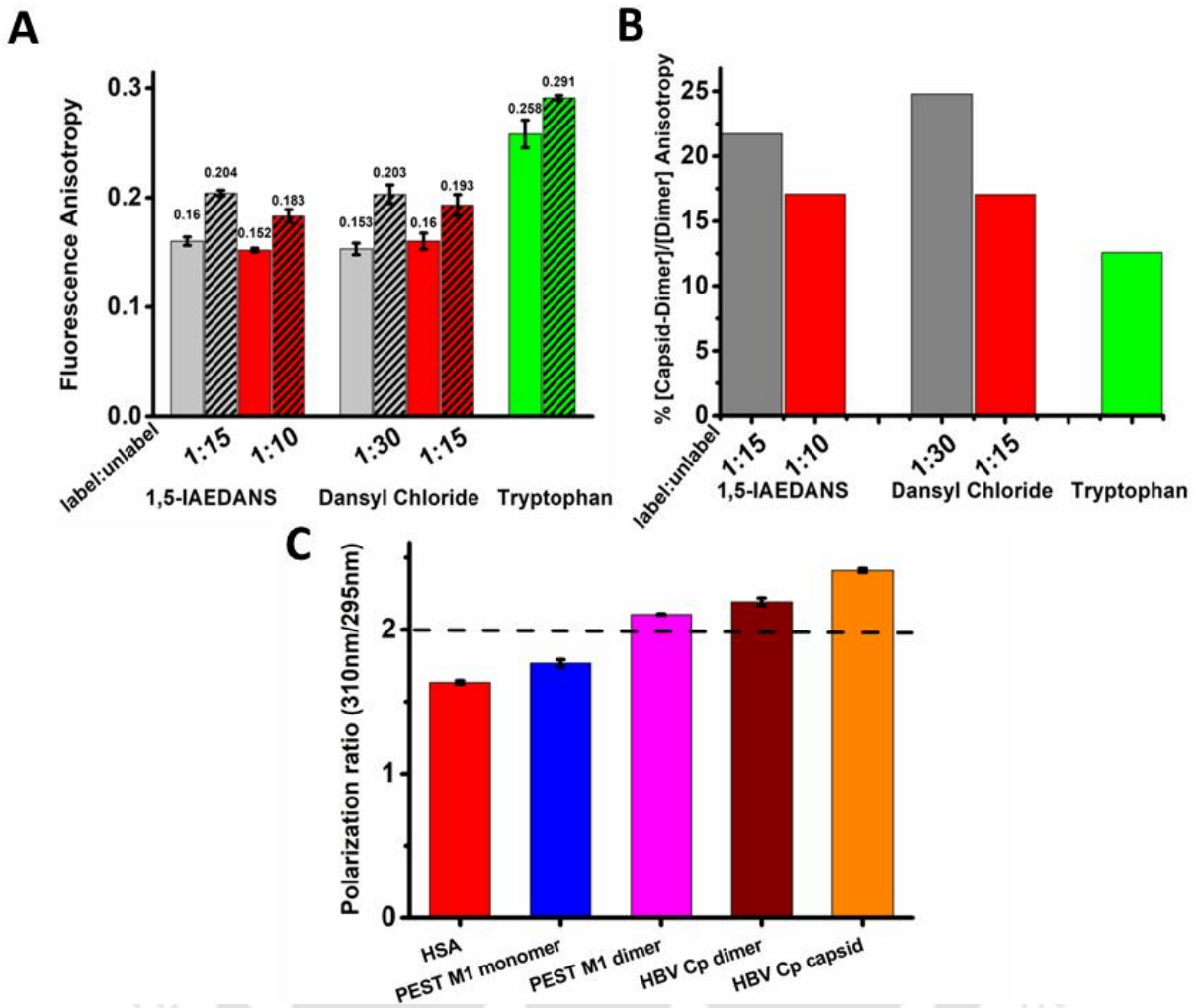


Figure 6.18: Steady-state fluorescence anisotropy for 1,5-IAEDANS, Dansyl Chloride attached to HBV Cp dimer and Tryptophan, is shown in subfigure A. Shaded represents capsids, whereas non-shaded represents dimer. Subfigure B shows the % change in fluorescence anisotropy in capsids compared to dimers. Subfigure C shows the occurrence of homo-FRET in HBV Cp dimer and capsids.

It was observed in **Figure 6.18A** that fluorescence anisotropy for both the extrinsic probes: 1,5-IAEDANS and Dansyl Chloride, as well as Tryptophan, increased considerably on capsid formation. The percentage change in fluorescence anisotropy upon capsid formation was calculated. It was found that the difference was more for the extrinsic probes than Tryptophan. Moreover, the changes observed for lower labelled/unlabelled dimer concentrations were marginally higher for both the probes. The percentage change in fluorescence anisotropy was most for low labelled dimers (> 20%), followed closely by the relatively highly labelled dimers (~ 17%) and the least for Tryptophan (~ 12%), as shown in **Figure 6.18B**.

The sensitivity of fluorescence anisotropy depends on the rotational correlation time of the macromolecule and the mean fluorescence lifetime of the tagged fluorophore. A fluorophore with a significantly shorter mean fluorescence than the rotational correlation time cannot reflect changes in fluorescence anisotropy. Both 1,5-IAEDANS and Dansyl Chloride generally show a relatively long mean fluorescence lifetime, around 10-15 ns, when attached to proteins[199]. Tryptophan, on the other hand, shows a mean fluorescence lifetime below 3.2 ns in proteins[2]. HBV capsids are large macromolecules with a mean diameter of 34 nm and long rotational correlation time. Thus, the labelled extrinsic probes with relatively longer mean lifetimes were more sensitive to the fluorescence anisotropy changes and could characterize the dimers from the capsids. In fact, the main motive behind labelling the dimers was to analyze the HBV capsid using such extrinsic probes. However, the subtle changes in Tryptophan fluorescence anisotropy were mainly because of the changes in the local conformation in its vicinity during HBV capsid formation. Also, the diminished sensitivity of Tryptophan fluorescence anisotropy to the slow tumbling of capsid can be attributed to the possibility of Trp-Trp homo-FRET. Decrease in Tryptophan fluorescence anisotropy has been observed in c-Myc PEST M1 dimers but not in monomers, due to possible homo-FRET phenomenon[200]. The Fὸster critical distance, R_0 value is 6-12 Å for a Trp-Trp pair [201] and thus, homo-FRET can be observed within distance of 2 R_0, i.e., 12-24 Å. **Figure 6.18C** shows homo-FRET in HBV Cp dimers which is more pronounced in capsids, indicated by the polarization ratio (310/295) greater than 2 in protein[202]. During HBV capsid assembly, the dimer-dimer contacts result in the two of the Trp residues (Trp102 and Trp125) in each monomeric unit to come within 13-18 Å, as obtained from PDB structure of HBV Cp149 (1QGT). Thus, the possible interference from Trp homo-FRET can contribute to the marginal increase in Trp fluorescence anisotropy.

6.2.9. Effect of 6 M Gdn.HCl on the steady-state fluorescence

In this section, the effect of Gdn.HCl on the structure of HBV dimers was monitored and subsequently compared to the capsids. In general, 6 M Gdn.HCl is often used to completely denature proteins and study protein folding and unfolding mechanisms[155]. Here, 6 M Gdn.HCl was used to denature HBV dimers completely, and the fluorescence intensity changes from the extrinsic probes: 1,5-IAEDANS and Tryptophan were analyzed. Furthermore, the changes associated with the fluorescence could give valuable insights into the position of both 1,5-IAEDANS and Tryptophan in the HBV dimers. Since both the 1,5-IAEDANS and Dansyl

Chloride showed similar behaviour with no changes in the steady state fluorescence upon capsid assembly, as seen in **Figures 6.16** and **6.17**. We studied the unfolding of the 1,5-IAEDANS labelled dimers and intrinsic Tryptophan fluorescence.

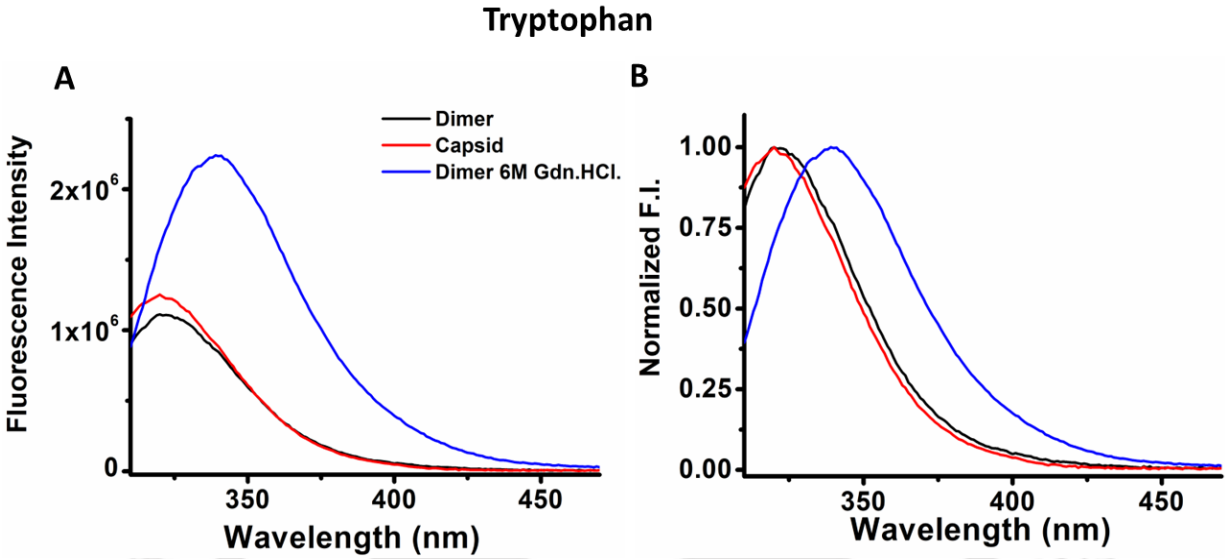


Figure 5.19: Steady-state Luminescence intensity(A) and Normalized Intensity (B) of Tryptophan in 6 M Gdn.HCl, Dimers and Capsid of HBV

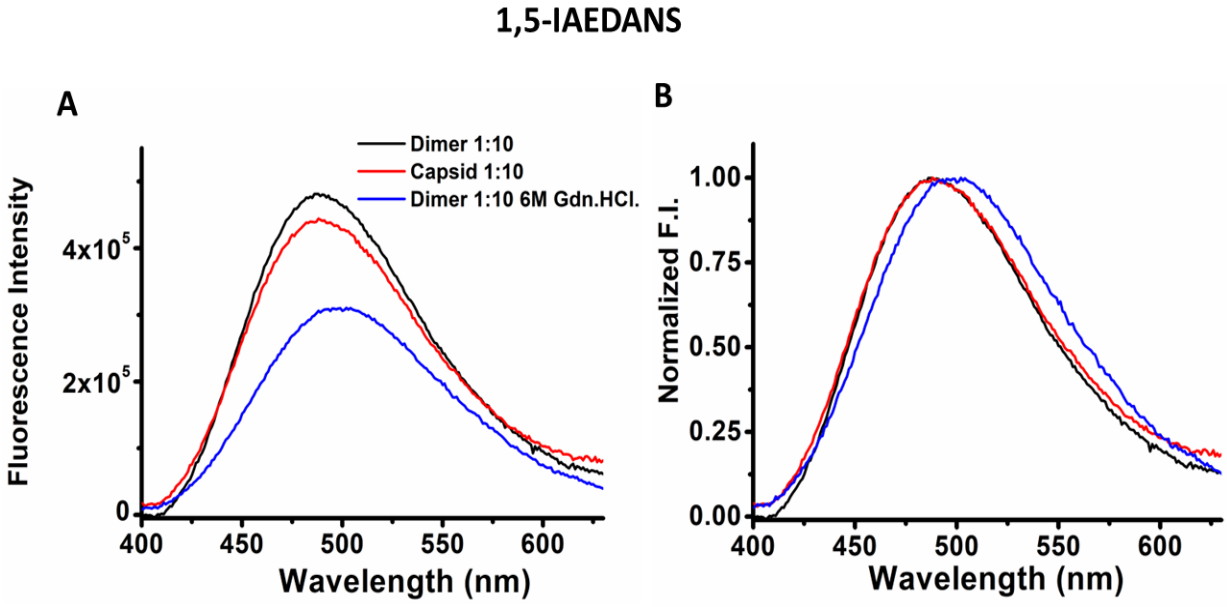


Figure 6.20: Steady-state Luminescence intensity (A) and Normalized Intensity (B) of 1,5-IAEDANS labelled in 6 M Gdn.HCL, Dimers and Capsid of HBV

Here, it could be seen in **Figure 6.19** that upon unfolding, the fluorescence emission maxima (λ_{max}) of Tryptophan increased from 323 nm to 339 nm during capsid formation. Interestingly,

the increase in the λ_{max} towards a longer wavelength is accompanied by enhanced luminescence intensity. It points out the effective quenching in most of the Tryptophan buried inside the HBV dimer. Upon unfolding, all the Tryptophan residues get exposed to solvent, thereby minimizing the quenching, which showed the red-shifted enhanced luminescence emission. Our result is in agreement with a similar observation from Tryptophan fluorescence in HBV dimers[150]. Similarly, for 1,5-IAEDANS labelled dimers, it could be seen in **Figure 6.20** that upon unfolding, the fluorescence emission maxima (λ_{max}) increased marginally from 488 nm to 502 nm during capsid formation. The increase in the λ_{max} towards a longer wavelength is accompanied by decreased luminescence intensity.

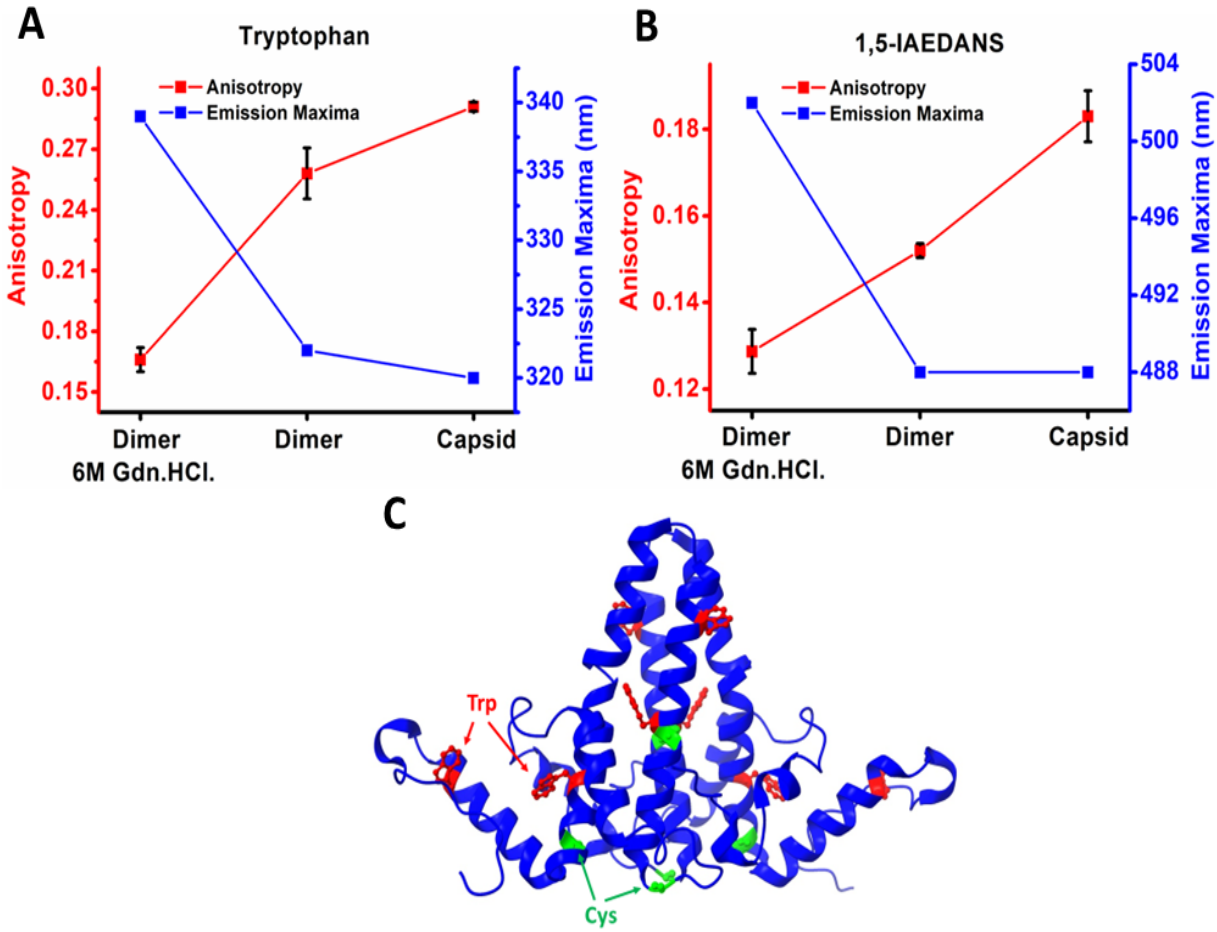


Figure 6.21: Steady-state Luminescence Anisotropy and Emission Maxima of Tryptophan (A) and labelled 1,5-IAEDANS (B) for 6 M Gdn.HCl denatured dimers, native dimers and capsid of HBV. Subfigure C shows the structure of the HBV dimer highlighting Tryptophan (red) and Cysteine (green).

To better understand the steady state fluorescence of 1,5-IAEDANS and Tryptophan in the HBV dimers, the changes observed in the fluorescence anisotropy and λ_{max} were plotted for

6M Gdn.HCl unfolded dimers, native dimers and capsids. As can be seen in **Figures 6.21A** and **B**, the fluorescence anisotropy and λ_{max} were highly anti-correlated. As the HBV dimers approached a more compact structure, the fluorescence anisotropy increased with a decrease in λ_{max}. It should be noted although the change λ_{max} is negligible for 1,5-IAEDANS from dimers to capsid, the slope observed for fluorescence anisotropy essentially increased. However, the slope for fluorescence anisotropy for Tryptophan from dimers to capsid flattened out compared to that from 6 M Gdn.HCl unfolded dimers to native dimers. At this point, it can be justified that the sensitivity in the fluorescence anisotropy for 1,5-IAEDANS labelled dimers is better in distinguishing HBV capsids from the dimers than intrinsic fluorescence from Tryptophan. We can thus conclude that during capsid assembly, the Tryptophan (highlighted in red) make weak dimer-dimer contacts and thus, the blue shift in λ_{max} was observed for Tryptophan. However, the free cysteines in the dimer are positioned in such a way that it remains solvent exposed irrespective of whether in dimer or in capsids as shown in **Figure 6.21C**. Thus, no significant change in the local environment surrounding the labelled 1,5-IAEDANS is expected. As a result, the λ_{max} observed for 1,5-IAEDANS showed no change upon assembly formation.

On the other hand, the increase in the fluorescence anisotropy in 1,5-IAEDANS is highly dependent on the increased rotational correlation time upon capsid formation and is not much dependent on the local environment and position of the free cysteines. Thus, the changes in fluorescence anisotropy during capsid assembly are more pronounced in 1,5-IAEDANS labelled dimers. However, the marginal increase in Tryptophan fluorescence anisotropy mainly arises from the changes in the local conformation and not primarily due to the addition of HBV dimers during HBV capsid formation.

In this study, only a tiny fraction of labelled dimers was required for the fluorescence anisotropy study as opposed to already established quenching experiments where higher labelled: un-labelled fractions were used. Fluorescence quenching of BODIPY dye in capsids has been used to monitor HBV capsid assembly in real time, where a labelling ratio of 1:4 was used[203]. Such high labelling ratio necessitates the need for more HBV dimers to be labelled for each set of reactions. Thus, it can be said that the steady-state fluorescence anisotropy can prove to be a promising tool for characterizing HBV capsids.

6.3. Conclusions

- HBV Cp dimer was successfully purified and assembled under 50 mM HEPES plus 0.25-0.5 M NaCl (pH 7.5). Subsequently, the dimers and capsids have been characterized by conventional methods: CD, DLS, AFM, TEM and SEC.

- Capsid assembly kinetics by 90° light scattering further confirmed the successful capsid assembly. The $t_{1/2}$ was less than 2 minutes at both 320 nm and 340 nm excitation.

- HBV capsid formation was monitored in real-time by both ProCharTS absorbance and luminescence. The kinetics could pick up the relatively later stages in capsid assembly at different excitation wavelengths with $t_{1/2}$ ~3-5 minutes.

- HBV dimers were successfully labelled with both Dansyl Chloride and 1,5-IAEDANS. Compared to dimers, enhanced fluorescence anisotropy (~25%) was observed upon capsid formation.

- Thus, ProCharTS can be a promising tool to study the effect of metal ions and RNA interaction during capsid formation and for rapid screening of antiviral drugs, i.e., capsid allosteric modulators (CpAMs).

Chapter 7

Thesis Summary and Future Directions

7.1. Thesis summary

Novel intrinsic chromophores in apo-proteins lacking aromatic amino acids has gained importance, and understanding the underlying phenomenon behind such observation has been a recent topic of interest. Here in this thesis, ProCharTS, which is proposed to be the origin of such intrinsic absorbance and luminescence features, is being studied. Although the ProCharTS phenomenon has been observed recently in highly charged Lysine-rich proteins, and the major contribution of Lysine has already been ascertained. Here we report for the first time ProCharTS in proteins and peptides rich in Arginine, Aspartate residues lacking Lysine residues. Similar absorbance and luminescence properties as that in Lysine-rich protein are observed. ProCharTS is found to be independent on the nature or type of charged amino acids and is a highlighting aspect of this study. Thus, ProCharTS can be expected to be evident as long as charge transfer takes place, irrespective of the type of charged species. Since ProCharTS and conventional chromophores like Tryptophan share a similar spectral domain in the UV-Visible region, the influence of ProCharTS on the indole fluorescence in Tryptophan is studied. The luminescence from the charged amino acids is found to contaminate the indole fluorescence with a relative broader luminescence spectrum, accompanied with a slight red-shift in the longer wavelength region. One exciting observation is the propensity of these charged residues to sully the indole fluorescence, irrespective of whether its fluorescence is getting quenched or not. Moreover, the decay kinetics from the excited state population in the indole ring is found to be affected by the presence of charged residues, thereby generating multiple luminescence lifetimes. This result could further explain the multi-exponentiality of Tryptophan often observed in proteins.

Finally, we focussed our attention on the applicability of ProCharTS on peptide aggregation and capsid assembly. It is found that ProCharTS is able to monitor the aggregation of Aβ peptides and HBV Cp 149 capsid assembly in *real-time*. The interconversion of the random coil to β-sheets in these Aβ peptides observed in CD highly correlated with our novel ProCharTS absorbance. Similarly, ProCharTS could detect the formation of the large clusters of 120 HBV Cp dimers arranged in an icosahedral geometry in capsids. During aggregation or capsid assembly, the increased interactions among the charged residues at close proximity in their oligomeric states or capsids are proposed to enhance ProCharTS signal. ProCharTS, being a simple technique, can thus be used for rapid initial screening of drugs for amyloid-linked diseases and viral infections. Moreover, this novel ProCharTS is a label-free technique which

does not interfere with the aggregation or capsid assembly kinetics. Additionally, the HBV capsids has been characterized using the fluorescence anisotropy from Dansyl Chloride and 1,5-IAEDANS probes attached to HBV Cp 149 dimers. Our approach uses relatively lower labelling ratio of fluorescence dyes. It thus eliminates the need for more HBV dimers to be labelled as opposed to already established quenching experiments, where higher labelled: un-labelled fractions were used.

7.2. Future directions

The ProCharTS phenomenon has been observed in monomeric proteins as well as in aggregates, and the major contribution of extinction coefficient for the observed ProCharTS phenomenon has already been established. But one inherent limitation that restricts the use of ProCharTS from mere spectroscopy to *in vivo* studies is the significantly low quantum yield. Thus, effectively modulating the quantum efficiency in ProCharTS by harnessing these electrons to recombine radiatively would help develop ProCharTS as a better tool for studying different protein dynamics. Moreover, there is a need to investigate the effect of post-translational modification, such as phosphorylation and acetylation in proteins. Phosphorylation increases the negative charge content, whereas acetylation results in decline in the positive charge in proteins. Thus, it would be interesting to look for whether ProCharTS can pick the subtle changes in the electron distribution in the protein. Also, as the kinetics for both the aggregation of Aβ peptides and truncated HBV Cp 149 capsid assembly in *real-time* was monitored using ProCharTS, it will be interesting to investigate the effect of potential drugs that inhibit or accelerate the kinetics. Screening of potential drugs based on ProCharTS would widen the applicability of ProCharTS even further. Moreover, full-length HBV Cp 183 is known to encapsidate heterologous RNA inside the capsid. The strong electrostatic interaction of the highly positively charged C-terminal domain in the full-length HBV Cp 183 with the nucleic acid is responsible for encapsulating the genetic material. Thus, the study of the assembly kinetics of the nucleic-acid-filled capsids using ProCharTS would be an interesting topic to work on. In fact, the Electron Transfer during the interaction of nucleic acid with any highly charged rich protein would provide a reliable platform to understand the origin of ProCharTS even further.

List of Publications

Published Journal(s)

Ansari MZ, Alom SE, Swaminathan R. Ordered structure induced in human c-Myc PEST region upon forming a disulphide bonded dimer J. Chem. Sci. 2021;133:26

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

Alom SE, Swaminathan R. ProCharTS arising from Arginine, Aspartate amino acids observed in peptides & proteins which are devoid of Lysine residues *Phys. Chem. Chem. Phys.* (*communicated*)

Alom SE, Kawa AH, Kalita S, Mandal B, Swaminathan R. ProCharTS to characterize the early events of aggregation of $A\beta_{16-22}$-derived switch-peptides (*Manuscript under preparation*)

Alom SE, Swaminathan K, Nuzelu V., de Rocquigny H, Swaminathan R. A rapid label-free approach monitoring Hepatitis B Virus Core protein (HBV Cp) capsid assembly in real-time: ProCharTS (*Manuscript under preparation*)

Conference Publication(s)

Alom SE, de Rocquigny H, Swaminathan R. Monitoring the hepatitis B viral (HBV) capsid assembly by charge transfer spectroscopy. Biophysical Journal. 2022;121(3):332a-3a.

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

List of Conferences attended

Conference(s)

Participated and presented a talk in "Light-matter interaction from scratch: Theory and Experiments at the border with biology" (ICTP 2021), Nov 22-25, organized by International Centre for Theoretical Physics (ICTP), Italy

Participated and presented poster online in the international conference "66th Annual Meeting of Biophysical Society" (BPS2022), Feb 19-23, organized by BPS in San Francisco, California.

Presented poster in international conference "64th Annual Meeting of Biophysical Society" (BPS2020), Feb 15-19, organized by BPS in San Diego, California.

Participated in "National Workshop on Fluorescence and Raman Spectroscopy" (FCS2018), Nov 12-17, organized by JNU and IIT Delhi, India

Participated and volunteered in "National Workshop on Fluorescence and Raman Spectroscopy" (FCS2017), Dec 17-21, organized by IIT Guwahati, India

Participated and presented a poster in 1st Departmental Retreat, organized by Dept of BSBE, IIT Guwahati, 21st Dec 2019

References

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

2. Lakowicz, J. R. Principles of fluorescence spectroscopy. *Princ. Fluoresc. Spectrosc.* 1–954 (2006)

3. Quickenden, T. I. & Irvin, J. A. The ultraviolet absorption spectrum of liquid water. *J. Chem. Phys.* **72**, 4416–4428 (1979).

4. Serdyuk, I. N., Zaccai, N. R. & Zaccai, J. Methods in Molecular Biophysics. *Methods Mol. Biophys.* (2007) doi:10.1017/cbo9780511811166.

5. Leermakers, P. A. & Vesley, G. F. Organic photochemistry and the excited state. *J. Chem. Educ.* **41**, 535 (1964).

6. Nilapwar, S. M., Nardelli, M., Westerhoff, H. V & Verma, M. Absorption spectroscopy. *Methods Enzymol.* **500**, 59–75 (2011).

7. Ham, J. S. & Platt, J. R. Far U.V. Spectra of Peptides. *J. Chem. Phys.* **20**, 335–336 (1952).

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

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

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

11. Antosiewicz, J. M. & Shugar, D. UV-Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 2: selected applications. *Biophys. Rev.* **8**, 163–177 (2016).

12. Bent, D. V & Hayon, E. Excited state chemistry of aromatic amino acids and related peptides. II. Phenylalanine. *J. Am. Chem. Soc.* **97**, 2606–2612 (1975).

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

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

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

16. Campagna, S., Puntoriero, F., Nastasi, F., Bergamini, G. & Balzani, V. Photochemistry and Photophysics of Coordination Compounds: Ruthenium *Photochem. photophysics Coord. Compd. I* 117–214 (2007).

17. Crosby, G. A. Spectroscopic investigations of excited states of transition-metal complexes. *Acc. Chem. Res.* **8**, 231–238 (1975).

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

19. Polo, F., Antonello, S., Formaggio, F., Toniolo, C. & Maran, F. Evidence against the hopping mechanism as an important electron transfer pathway for conformationally constrained oligopeptides. *J. Am. Chem. Soc.* **127**, 492–493 (2005).

20. Jortner, J., Bixon, M., Langenbacher, T. & Michel-Beyerle, M. E. Charge transfer and transport in DNA. *Proc. Natl. Acad. Sci.* **95**, 12759–12765 (1998).

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

22. Kai, M., Takeda, K., Morita, T. & Kimura, S. Distance dependence of long-range electron transfer through helical peptides. *J. Pept. Sci. an Off. Publ. Eur. Pept. Soc.* **14**, 192–202 (2008).

23. Prasad, S. *et al.* Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. *Chem. Sci.* **8**, 5416–5433 (2017).

24. Ryzhkina, I. S. *et al.* Self-organization and properties of dispersed systems based on dilute aqueous solutions of (S)- and (R)-lysine. *Mendeleev Commun.* **28**, 66–69 (2018).

25. Homchaudhuri, L. & Swaminathan, R. Novel Absorption and Fluorescence Characteristics of L-Lysine. *Chem. Lett.* **2001**, 844–845 (2001).

26. 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).

27. Chen, X. *et al.* Prevalent intrinsic emission from nonaromatic amino acids and poly(amino acids). *Sci. China Chem.* **61**, 351–359 (2018).

28. Stephens, A. D. *et al.* Intrinsic fluorescence in non-aromatic peptide structures is induced by collective vibrations, charge reorganisation and short hydrogen bonds, as shown in a new glutamine-related structure. *bioRxiv* 2001–2020 (2020).

29. Chai, B. H., Zheng, J. M., Zhao, Q. & Pollack, G. H. Spectroscopic studies of solutes in aqueous solution. *J. Phys. Chem. A* **112**, 2242–2247 (2008).

30. 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).

31. Guptasarma, P. Solution-state characteristics of the ultraviolet A-induced visible fluorescence from proteins. *Arch. Biochem. Biophys.* **478**, 127–129 (2008).

32. Bhattacharya, A. *et al.* Direct Evidence of Intrinsic Blue Fluorescence from Oligomeric Interfaces of Human Serum Albumin. *Langmuir* **33**, 10606–10615 (2017).

33. 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).

34. Chan, F. T. S. *et al.* Protein amyloids develop an intrinsic fluorescence signature during aggregation. *Analyst* **138**, 2156–2162 (2013).

35. Tikhonova, T. N. *et al.* Dissection of the deep-blue autofluorescence changes accompanying amyloid fibrillation. *Arch. Biochem. Biophys.* **651**, 13–20 (2018).

36. Konopka, M. *et al.* Determination of non-traditional intrinsic fluorescence (NTIF) emission sites in 1-(4-carbomethoxypyrrolidone)-PAMAM dendrimers using CNDP-based quenching studies. *J. Nanoparticle Res.* **20**, 1–11 (2018).

37. Zheng, H., Cao, P., Wang, Y., Lu, X. & Wu, P. Ultralong Room-Temperature Phosphorescence from Boric Acid. *Angew. Chemie* **133**, 9586–9592 (2021).

38. Zhu, T., Yang, T., Zhang, Q. & Yuan, W. Z. Clustering and halogen effects enabled red/near-infrared room temperature phosphorescence from aliphatic cyclic imides. *Nat. Commun.* **13**, 1–12 (2022).

39. Zhang Yuan, W. & Zhang, Y. Nonconventional macromolecular luminogens with aggregation-induced emission characteristics. *J. Polym. Sci. Part A Polym. Chem.* **55**, 560–574 (2017).

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. 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).

42. Mandal, I., Manna, S. & Venkatramani, R. UV-Visible Lysine-Glutamate Dimer Excitations in Protein Charge Transfer Spectra: TDDFT Descriptions Using an Optimally Tuned CAM-B3LYP Functional. *J. Phys. Chem. B* **123**, 10967–10979 (2019).

43. 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).

44. Pinotsi, D. *et al.* Proton Transfer and Structure-Specific Fluorescence in Hydrogen Bond-Rich Protein Structures. *J. Am. Chem. Soc.* **138**, 3046–3057 (2016).

45. 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).

46. Xu, Z. *et al.* Crystal structure of D-Hydantoinase from Burkholderia pickettii at a resolution of 2.7 Angstroms: insights into the molecular basis of enzyme thermostability. *J. Bacteriol.* **185**, 4038–4049 (2003).

47. Bhattacharya, A., Prajapati, R., Chatterjee, S. & Mukherjee, T. K. Concentration-dependent reversible self-oligomerization of serum albumins through intermolecular β-sheet formation. *Langmuir* **30**, 14894–14904 (2014).

48. Stagi, L., Malfatti, L., Caboi, F. & Innocenzi, P. Thermal Induced Polymerization of l-Lysine forms Branched Particles with Blue Fluorescence. *Macromol. Chem. Phys.* **222**, (2021).

49. 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).

50. Xue, C., Lin, T. Y., Chang, D. & Guo, Z. Thioflavin T as an amyloid dye: Fibril quantification, optimal concentration and effect on aggregation. *R. Soc. Open Sci.* **4**, (2017).

51. Hudson, S. A., Ecroyd, H., Kee, T. W. & Carver, J. A. The thioflavin T fluorescence assay for amyloid fibril detection can be biased by the presence of exogenous compounds. *FEBS J.* **276**, 5960–5972 (2009).

52. Szabo, A. G. & Rayner, D. M. Fluorescence Decay of Tryptophan Conformers in Aqueous Solution. *J. Am. Chem. Soc.* **102**, 554–563 (1980).

53. Clayton, A. H. A. & Sawyer, W. H. Tryptophan rotamer distributions in amphipathic peptides at a lipid surface. *Biophys. J.* **76**, 3235–3242 (1999).

54. Dahms, T. E., Dahms, T. E., Szabo, A. G., Willis, K. J. & Szabo, A. G. Conformational Heterogeneity of Tryptophan in a Protein Crystal. *J. Am. Chem. Soc.* **117**, 2321–2326 (1995).

55. Ghisaidoobe, A. B. T. & Chung, S. J. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. *Int. J. Mol. Sci.* **15**, 22518–22538 (2014).

56. Pan, C.-P. & Barkley, M. D. Conformational Effects on Tryptophan Fluorescence in Cyclic Hexapeptides. *Biophys. J.* **86**, 3828–3835 (2004).

57. Grinvald, A. & Steinberg, I. Z. Fast Relaxation Processes in a Protein Revealed by the Decay Kinetics of Tryptophan Fluorescence. *Biochemistry* **13**, 5170–5178 (1974).

58. Hudson, B. S., Huston, J. M. & Soto-Campos, G. A Reversible "Dark State" Mechanism for Complexity of the Fluorescence of Tryptophan in Proteins. *J. Phys. Chem. A* **103**, 2227–2234 (1999).

59. Cowgill, R. W. Fluorescence and the structure of proteins. I. Effects of substituents on the fluorescence of indole and phenol compounds. *Arch. Biochem. Biophys.* **100**, 36–44 (1963).

60. Liu, T. *et al.* Ionization potentials of fluoroindoles and the origin of nonexponential tryptophan fluorescence decay in proteins. *J. Am. Chem. Soc.* **127**, 4104–4113 (2005).

61. De Beuckeleer, K., Volckaert, G. & Engelborghs, Y. Time resolved fluorescence and phosphorescence properties of the individual tryptophan residues of barnase: Evidence for protein-protein interactions. *Proteins Struct. Funct. Genet.* **36**, 42–53 (1999).

62. Weisenborn, P. C. M., Meder, H., Egmond, M. R., Visser, T. J. W. G. & Van Hoek, A. Photophysics of the single tryptophan residue in Fusarium solani Cutinase: Evidence for the occurrence of conformational substates with unusual fluorescence behaviour. *Biophys. Chem.* **58**, 281–288 (1996).

63. Steiner, R. F. & Kirby, E. P. The interaction of the ground and excited states of indole derivatives with electron scavengers. *J. Phys. Chem.* **73**, 4130–4135 (1969).

64. Yu, H. T., Colucci, W. J., McLaughlin, M. L. & Barkley, M. D. Fluorescence quenching in indoles by excited-state proton transfer. *J. Am. Chem. Soc.* **114**, 8455–8463 (1992).

65. Chen, Y. & Barkley, M. D. Toward understanding tryptophan fluorescence in proteins. *Biochemistry* **37**, 9976–9982 (1998).

66. Vivian, J. T. & Callis, P. R. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.* **80**, 2093–2109 (2001).

67. 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).

68. Prince, P. M., Ali, G. & Ali, G. World Alzheimer Report 2015 The Global Impact of Dementia. (2015).

69. Wu, K. Y., Doan, D., Medrano, M. & Chang, C. A. Modeling structural interconversion in Alzheimers' amyloid beta peptide with classical and intrinsically disordered protein force fields. *J. Biomol. Struct. Dyn.* 1–18 (2021).

70. Soto, C., Brañes, M. C., Alvarez, J. & Inestrosa, N. C. Structural determinants of the Alzheimer's amyloid beta-peptide. *J. Neurochem.* **63**, 1191–1198 (1994).

71. Kirschner, D. A. *et al.* Synthetic peptide homologous to beta protein from Alzheimer disease forms amyloid-like fibrils in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6953–6957 (1987).

72. McLean, C. A. *et al.* Soluble pool of Aβ amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. *Ann. Neurol.* **46**, 860–866 (1999).

73. Rosenblum, W. I. Structure and location of amyloid beta peptide chains and arrays in Alzheimer's disease: new findings require reevaluation of the amyloid hypothesis and of tests of the hypothesis. *Neurobiol. Aging* **23**, 225–230 (2002).

74. Walsh, D. M. *et al.* Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* **416**, 535–539 (2002).

75. Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat. Rev. Mol. Cell Biol.* **8**, 101–112 (2007).

76. Viola, K. L. *et al.* Towards non-invasive diagnostic imaging of early-stage Alzheimer's disease. *Nat. Nanotechnol.* **10**, 91–98 (2015).

77. Colletier, J. P. *et al.* Molecular basis for amyloid-β polymorphism. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16938–16943 (2011).

78. Petkova, A. T. *et al.* A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 16742–16747 (2002).

79. Paravastu, A. K., Leapman, R. D., Yau, W. M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer's β-amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18349–18354 (2008).

80. Van Nostrand, W. E., Melchor, J. P., Cho, H. S., Greenberg, S. M. & Rebeck, G. W.

Pathogenic effects of D23N Iowa mutant amyloid beta -protein. *J. Biol. Chem.* **276**, 32860–32866 (2001).

81. Tycko, R., Sciarretta, K. L., Orgel, J. P. R. O. & Meredith, S. C. Evidence for novel beta-sheet structures in Iowa mutant beta-amyloid fibrils. *Biochemistry* **48**, 6072–6084 (2009).

82. Lansbury, P. T. & Lashuel, H. A. A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature* **443**, 774–779 (2006).

83. Mutter, M. *et al.* Switch peptides in statu nascendi: Induction of conformational transitions relevant to degenerative diseases. *Angew. Chemie - Int. Ed.* **43**, 4172–4178 (2004).

84. Dos Santos, S. *et al.* Switch-peptides: Controlling self-assembly of amyloid β-derived peptides in vitro by consecutive triggering of acyl migrations. *J. Am. Chem. Soc.* **127**, 11888–11889 (2005).

85. Giamblanco, N. *et al.* Detection of protein aggregate morphology through single antifouling nanopore. *Sensors Actuators B Chem.* **260**, (2018).

86. Lashuel, H. A. *et al.* Mixtures of wild-type and a pathogenic (E22G) form of Aβ40 in vitro accumulate protofibrils, including amyloid pores. *J. Mol. Biol.* **332**, 795–808 (2003).

87. Lee, C. C., Nayak, A., Sethuraman, A., Belfort, G. & McRae, G. J. A three-stage kinetic model of amyloid fibrillation. *Biophys. J.* **92**, 3448–3458 (2007).

88. Janek, K. *et al.* Study of the conformational transition of Aβ(1-42) using D-amino acid replacement analogues. *Biochemistry* **40**, 5457–5463 (2001).

89. Wolff, M. *et al.* Amyloid β oligomeric species present in the lag phase of amyloid formation. *PLoS One* **10**, 1–14 (2015).

90. Sunde, M. *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273**, 729–739 (1997).

91. Inouye, H., Fraser, P. E. & Kirschner, D. A. Structure of beta-crystallite assemblies formed by Alzheimer beta-amyloid protein analogues: analysis by x-ray diffraction. *Biophys. J.* **64**, 502–519 (1993).

92. Ahmed, M. *et al.* Structural conversion of neurotoxic amyloid-beta(1-42) oligomers to fibrils. *Nat. Struct. & Mol. Biol.* **17**, 561–567 (2010).

93. Zhang, Y. *et al.* Monitoring amyloid-B proteins aggregation based on label-free aptasensor. *Sensors Actuators, B Chem.* **288**, 535–542 (2019).

94. Sharpe, S., Simonetti, K., Yau, J. & Walsh, P. Solid-State NMR characterization of autofluorescent fibrils formed by the elastin-derived peptide GVGVAGVG. *Biomacromolecules* **12**, 1546–1555 (2011).

95. Jiang, D., Rauda, I., Han, S., Chen, S. & Zhou, F. Aggregation Pathways of the Amyloid β(1–42) Peptide Depend on Its Colloidal Stability and Ordered β-Sheet Stacking. *Langmuir* **28**, 12711–12721 (2012).

96. Kirschner, D. A., Abraham, C. & Selkoe, D. J. X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer disease indicates cross-beta conformation. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 503–507 (1986).

97. Ganem, D. Assembly of hepadnaviral virions and subviral particles. *Curr. Top. Microbiol. Immunol.* **168**, 61–83 (1991).

98. Dane, D. S., Cameron, C. H. & Briggs, M. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet (London, England)* **1**, 695–698 (1970).

99. Robinson, W. S. & Lutwick, L. I. The virus of hepatitis, type B (first of two parts). *N. Engl. J. Med.* **295**, 1168–1175 (1976).

100. Summers, J., O'Connell, A. & Millman, I. Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 4597–4601 (1975).

101. Beck, J. & Nassal, M. Hepatitis B virus replication. *World J. Gastroenterol.* **13**, 48–64 (2007).

102. Nassal, M. Hepatitis B viruses: reverse transcription a different way. *Virus Res.* **134**, 235–249 (2008).

103. Jones, S. A. & Hu, J. Hepatitis B virus reverse transcriptase: diverse functions as classical and emerging targets for antiviral intervention. *Emerg. Microbes Infect.* **2**, e56 (2013).

104. Tang, H., Oishi, N., Kaneko, S. & Murakami, S. Molecular functions and biological roles of hepatitis B virus x protein. *Cancer Sci.* **97**, 977–983 (2006).

105. Bock, C. T., Schwinn, S., Schröder, C. H., Velhagen, I. & Zentgraf, H. Localization of hepatitis B virus core protein and viral DNA at the nuclear membrane. *Virus Genes* **12**, 53–63 (1996).

106. Lewellyn, E. B. & Loeb, D. D. The arginine clusters of the carboxy-terminal domain of the core protein of hepatitis B virus make pleiotropic contributions to genome replication. *J. Virol.* **85**, 1298–1309 (2011).

107. Scaglioni, P. P., Melegari, M. & Wands, J. R. Posttranscriptional regulation of hepatitis B virus replication by the precore protein. *J. Virol.* **71**, 345–353 (1997).

108. Guo, Y.-H., Li, Y.-N., Zhao, J.-R., Zhang, J. & Yan, Z. HBc binds to the CpG islands of HBV cccDNA and promotes an epigenetic permissive state. *Epigenetics* **6**, 720–726 (2011).

109. Gish, R. G. *et al.* Entecavir therapy for up to 96 weeks in patients with HBeAg-positive chronic hepatitis B. *Gastroenterology* **133**, 1437–1444 (2007).

110. Bourne, C. *et al.* Small-molecule effectors of hepatitis B virus capsid assembly give insight into virus life cycle. *J. Virol.* **82**, 10262–10270 (2008).

111. Belloni, L. *et al.* HAPs hepatitis B virus (HBV) capsid inhibitors prevent HBc interaction with the viral minichromosome and selected host cell genes to inhibits transcription and affect cccDNA stability. *Dig. Liver Dis.* **46**, e9 (2014).

112. Milich, D. R. *et al.* Comparative immunogenicity of hepatitis B virus core and E antigens. *J. Immunol.* **141**, 3617–3624 (1988).

113. Dandri, M., Volz, T. K., Lütgehetmann, M. & Petersen, J. Animal models for the study of HBV replication and its variants. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* **34 Suppl 1**, S54-62 (2005).

114. Stannard, L. M. & Hodgkiss, M. Morphological irregularities in Dane particle cores. *J. Gen. Virol.* **45**, 509–514 (1979).

115. Ning, X. *et al.* Secretion of genome-free hepatitis B virus--single strand blocking model for virion morphogenesis of para-retrovirus. *PLoS Pathog.* **7**, e1002255 (2011).

116. Chain, B. M. & Myers, R. Variability and conservation in hepatitis B virus core protein. *BMC Microbiol.* **5**, 33 (2005).

117. Wynne, S. A., Crowther, R. A. & Leslie, A. G. The crystal structure of the human hepatitis B virus capsid. *Mol. Cell* **3**, 771–780 (1999).

118. Chu, T.-H. *et al.* Nucleic Acid Chaperone Activity Associated with the Arginine-Rich Domain of Human Hepatitis B Virus Core Protein. *J. Virol.* **88**, 2530–2543 (2014).

119. Zhou, S. & Standring, D. N. Hepatitis B virus capsid particles are assembled from core-protein dimer precursors. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10046–10050 (1992).

120. Zheng, J., Schödel, F. & Peterson, D. L. The structure of hepadnaviral core antigens. Identification of free thiols and determination of the disulfide bonding pattern. *J. Biol. Chem.* **267**, 9422–9429 (1992).

121. Roseman, A. M., Berriman, J. A., Wynne, S. A., Butler, P. J. G. & Crowther, R. A. A structural model for maturation of the hepatitis B virus core. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15821–15826 (2005).

122. Porterfield, J. Z. *et al.* Full-Length Hepatitis B Virus Core Protein Packages Viral and Heterologous RNA with Similarly High Levels of Cooperativity. *J. Virol.* **84**, 7174–7184 (2010).

123. Ceres, P. & Zlotnick, A. Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry* **41**, 11525–11531 (2002).

124. Zlotnick, A., Ceres, P., Singh, S. & Johnson, J. M. A Small Molecule Inhibits and Misdirects Assembly of Hepatitis B Virus Capsids. *J. Virol.* **76**, 4848–4854 (2002).

125. Lopez, M. *et al.* Characterization of the size distribution and aggregation of virus-like nanoparticles used as active ingredients of the HeberNasvac therapeutic vaccine against chronic hepatitis B. *Adv. Nat. Sci. Nanosci. Nanotechnol.* **8**, (2017).

126. Uetrecht, C., Barbu, I. M., Shoemaker, G. K., Van Duijn, E. & Heck, A. J. R. Interrogating viral capsid assembly with ion mobility-mass spectrometry. *Nat. Chem.* **3**, 126–132 (2011).

127. Pierson, E. E. *et al.* Detection of late intermediates in virus capsid assembly by charge detection mass spectrometry. *J. Am. Chem. Soc.* **136**, 3536–3541 (2014).

128. Zhou, J. *et al.* Characterization of Virus Capsids and Their Assembly Intermediates by

Multicycle Resistive-Pulse Sensing with Four Pores in Series. *Anal. Chem.* **90**, 7267–7274 (2018).

129. Uetrecht, C. *et al.* High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9216–9220 (2008).

130. Schlicksup, C. J. *et al.* Hepatitis B virus core protein allosteric modulators can distort and disrupt intact capsids. *Elife* **7**, e31473 (2018).

131. Steppert, P. *et al.* Quantification and characterization of virus-like particles by size-exclusion chromatography and nanoparticle tracking analysis. *J. Chromatogr. A* **1487**, 89–99 (2017).

132. Kler, S. *et al.* RNA encapsidation by SV40-derived nanoparticles follows a rapid two-state mechanism. *J. Am. Chem. Soc.* **134**, 8823–8830 (2012).

133. Kramberger, P., Ciringer, M., Štrancar, A. & Peterka, M. Evaluation of nanoparticle tracking analysis for total virus particle determination. *Virol. J.* **9**, 265 (2012).

134. Han, Y. *et al.* Solid-State NMR Studies of HIV-1 Capsid Protein Assemblies. *J. Am. Chem. Soc.* **132**, 1976–1987 (2010).

135. Parveen, N., Borrenberghs, D., Rocha, S. & Hendrix, J. Single viruses on the fluorescence microscope: Imaging molecular mobility, interactions and structure sheds new light on viral replication. *Viruses* **10**, (2018).

136. Rayaprolu, V. *et al.* Comparative Analysis of Adeno-Associated Virus Capsid Stability and Dynamics. *J. Virol.* **87**, 13150–13160 (2013).

137. Cantor, C. R. & Schimmel, P. R. Techniques for the study of biological structure and function. (1980).

138. 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).

139. Swaminathan, R. & Periasamy, N. Analysis of fluorescence decay by the maximum entropy method: Influence of noise and analysis parameters on the width of the distribution of lifetimes. *Proc. Indian Acad. Sci. Chem. Sci.* **108**, 39–49 (1996).

140. Karnik, A. V & Hasan, M. *Stereochemistry: A Three-Dimensional Insight*. (Elsevier, 2021).

141. Van Holde, K. E., Johnson, W. C. & Ho, P. S. Principles of physical biochemistry. (2006).

142. Barron, A. R. Physical methods in chemistry and nano science. (2015).

143. Longo, L. M., Lee, J. & Blaber, M. Simplified protein design biased for prebiotic amino acids yields a foldable, halophilic protein. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2135–2139 (2013).

144. Zlotnick, A. *et al.* Dimorphism of Hepatitis B Virus Capsids Is Strongly Influenced by the C-Terminus of the Capsid Protein. **2960**, 7412–7421 (1996).

145. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual.*

(Cold spring harbor laboratory press, 1989).

146. 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).

147. Waddell, W. J. A simple ultraviolet spectrophotometric method for the determination of protein. *J. Lab. Clin. Med.* **48**, 311–314 (1956).

148. Sharma, B., Kalita, S., Paul, A., Mandal, B. & Paul, S. The role of caffeine as an inhibitor in the aggregation of amyloid forming peptides: A unified molecular dynamics simulation and experimental study. *RSC Adv.* **6**, 78548–78558 (2016).

149. Paul, A. *et al.* Amyloid β derived switch-peptides as a tool for investigation of early events of aggregation: A combined experimental and theoretical approach. *Medchemcomm* **7**, 311–316 (2016).

150. Wingfield, P. T., Stahl, S. J., Williams, R. W. & Steven1, A. C. Hepatitis core Antigen produced in Escherichia coli. *Biochemistry* **34**, 4919–4932 (1995).

151. Pfefferkorn, C. M. & Lee, J. C. 5-fluoro-D,L-tryptophan as a dual NMR and fluorescent probe of α-synuclein. *Methods Mol. Biol.* **895**, 197–209 (2012).

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

153. Hawe, A., Filipe, V. & Jiskoot, W. Fluorescent molecular rotors as dyes to characterize polysorbate-containing IgG formulations. *Pharm. Res.* **27**, 314–326 (2010).

154. Sulatskaya, A. I., Lavysh, A. V, Maskevich, A. A., Kuznetsova, I. M. & Turoverov, K. K. Thioflavin T fluoresces as excimer in highly concentrated aqueous solutions and as monomer being incorporated in amyloid fibrils. *Sci. Rep.* 7, 2146 (2017).

155. 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).

156. Bevington, P. R., Robinson, D. K., Blair, J. M., Mallinckrodt, A. J. & McKay, S. Data reduction and error analysis for the physical sciences. *Comput. Phys.* **7**, 415–416 (1993).

157. 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).

158. Chen, R. F. Dansyl labeled proteins: determination of extinction coefficienc and number of bound residues with radioactive dansyl chloride. *Anal. Biochem.* **25**, 412–416 (1968).

159. Levi, V. & González Flecha, F. L. Labeling of proteins with fluorescent probes: photophysical characterization of dansylated bovine serum albumin. *Biochem. Mol. Biol. Educ.* **31**, 333–336 (2003).

160. Wiederschain, G. Y. The Molecular Probes handbook. A guide to fluorescent probes and labeling technologies. *Biochem.* **76**, 1276–1277 (2011).

161. Hudson, E. N. & Weber, G. Synthesis and characterization of two fluorescent sulfhydryl reagents. *Biochemistry* **12**, 4154–4161 (1973).

162. Goldsmith, C. S. & Miller, S. E. Modern uses of electron microscopy for detection of viruses. *Clin. Microbiol. Rev.* **22**, 552–563 (2009).

163. Terpugov, E. L., Kondratyev, M. S. & Degtyareva, O. V. Light-induced effects in glycine aqueous solution studied by Fourier transform infrared-emission spectroscopy and ultraviolet-visible spectroscopy. *J. Biomol. Struct. Dyn.* **39**, 108–117 (2021).

164. Demchenko, A. P., Tomin, V. I. & Chou, P. T. Breaking the Kasha Rule for More Efficient Photochemistry. *Chem. Rev.* **117**, 13353–13381 (2017).

165. Scuppa, S., Orian, L., Donoli, A., Santi, S. & Meneghetti, M. Anti-Kasha's rule fluorescence emission in (2-ferrocenyl)indene generated by a twisted intramolecular charge-transfer (TICT) process. *J. Phys. Chem. A* **115**, 8344–8349 (2011).

166. Dou, W. T., Chen, W., He, X. P., Su, J. & Tian, H. Vibration-Induced-Emission (VIE) for imaging amyloid β fibrils. *Faraday Discuss.* **196**, 395–402 (2017).

167. Shen, X. Bin *et al.* Excitation-wavelength-dependent photoluminescence of silicon nanoparticles enabled by adjustment of surface ligands. *Chem. Commun.* **54**, 4947–4950 (2018).

168. Lai, Y. *et al.* Effective Internal and External Modulation of Nontraditional Intrinsic Luminescence. *Small* **16**, 1–6 (2020).

169. Arnon, Z. A. *et al.* On-off transition and ultrafast decay of amino acid luminescence driven by modulation of supramolecular packing. *iScience* **24**, 102695 (2021).

170. Ravanfar, R., Bayles, C. J. & Abbaspourrad, A. Structural Chemistry Enables Fluorescence of Amino Acids in the Crystalline Solid State. *Cryst. Growth Des.* **20**, 1673–1680 (2020).

171. Schmidt, N., Mishra, A., Lai, G. H. & Wong, G. C. L. Arginine-rich cell-penetrating peptides. *FEBS Lett.* **584**, 1806–1813 (2010).

172. Chen, Y., Liu, B., Yu, H.-T. & Barkley, M. D. The Peptide Bond Quenches Indole Fluorescence. *J. Am. Chem. Soc.* **118**, 9271–9278 (1996).

173. Juszczak, L. J. & Eisenberg, A. S. The Color of Cation-Ï€ Interactions: Subtleties of Amine-Tryptophan Interaction Energetics Allow for Radical-like Visible Absorbance and Fluorescence. *J. Am. Chem. Soc.* **139**, 8302–8311 (2017).

174. Scott, J. N. & Callis, P. R. Insensitivity of tryptophan fluorescence to local charge mutations. *J. Phys. Chem. B* **117**, 9598–9605 (2013).

175. Grinvald, A. & Steinberg, I. Z. On the analysis of fluorescence decay kinetics by the method of least-squares. *Anal. Biochem.* **59**, 583–598 (1974).

176. Jesus, C. S. H., Soares, H. T., Piedade, A. P., Cortes, L. & Serpa, C. Using amyloid autofluorescence as a biomarker for lysozyme aggregation inhibition. *Analyst* **146**, 2383–2391 (2021).

177. Fricano, A., Librizzi, F., Rao, E., Alfano, C. & Vetri, V. Blue autofluorescence in protein aggregates "lighted on" by UV induced oxidation. *Biochim. Biophys. Acta - Proteins Proteomics* **1867**, 140258 (2019).

178. Lindberg, D. J. *et al.* Binding of Thioflavin-T to Amyloid Fibrils Leads to Fluorescence Self-Quenching and Fibril Compaction. *Biochemistry* **56**, 2170–2174 (2017).

179. Singh, P., Wangoo, N. & Sharma, R. K. Phenylalanine dimer assembly structure as the basic building block of an amyloid like photoluminescent nanofibril network. *Soft Matter* **16**, 4105–4109 (2020).

180. Singh, P., Pandey, S. K., Grover, A., Sharma, R. K. & Wangoo, N. Understanding the self-ordering of amino acids into supramolecular architectures: Co-assembly-based modulation of phenylalanine nanofibrils. *Mater. Chem. Front.* **5**, 1971–1981 (2021).

181. Serrano-Andrés, L. & Fülscher, M. P. Charge transfer transitions in neutral and ionic polypeptides: A theoretical study. *J. Phys. Chem. B* **105**, 9323–9330 (2001).

182. del Mercato, L. L. *et al.* Charge transport and intrinsic fluorescence in amyloid-like fibrils. *Proc. Natl. Acad. Sci.* **104**, 18019–18024 (2007).

183. Apter, B. *et al.* Fluorescence waveguiding in amyloidogenic fibers. in *Neurophotonics* **11360**, 63–72 (SPIE, 2020).

184. Apter, B. *et al.* Fold-Sensitive Visible Fluorescence in β-Sheet Peptide Structures. *Adv. Opt. Mater.* **9**, 1–10 (2021).

185. Pansieri, J. *et al.* Ultraviolet–visible–near-infrared optical properties of amyloid fibrils shed light on amyloidogenesis. *Nat. Photonics* **13**, 473–479 (2019).

186. Monti, A. *et al.* Amyloid-like Prep1 peptides exhibit reversible blue-green-red fluorescencein vitroand in living cells. *Chem. Commun.* **57**, 3720–3723 (2021).

187. Ziaunys, M., Sneideris, T. & Smirnovas, V. Exploring the potential of deep-blue autofluorescence for monitoring amyloid fibril formation and dissociation. *PeerJ* **2019**, (2019).

188. Yu, J., Huang, D. M., Shapter, J. G. & Abell, A. D. Electrochemical and computational studies on intramolecular dissociative electron transfer in β-peptides. *J. Phys. Chem. C* **116**, 26608–26617 (2012).

189. Joseph, S. K. *et al.* Proton-Transfer-Induced Fluorescence in Self-Assembled Short Peptides. *J. Phys. Chem. A* **123**, 1758–1765 (2019).

190. Zlotnick, A. Are weak protein-protein interactions the general rule in capsid assembly? *Virology* **315**, 269–274 (2003).

191. Lutomski, C. A. *et al.* Multiple Pathways in Capsid Assembly. *J. Am. Chem. Soc.* **140**, 5784–5790 (2018).

192. Stray, S. J., Ceres, P. & Zlotnick, A. Zinc Ions Trigger Conformational Change and Oligomerization of Hepatitis B Virus Capsid Protein. *Biochemistry* **43**, 9989–9998 (2004).

193. Moisant, P., Neeman, H. & Zlotnick, A. Exploring the paths of (virus) assembly. *Biophys. J.* **99**, 1350–1357 (2010).

194. Asor, R., Schlicksup, C. J., Zhao, Z., Zlotnick, A. & Raviv, U. Rapidly Forming Early Intermediate Structures Dictate the Pathway of Capsid Assembly. *J. Am. Chem. Soc.*

142, 7868–7882 (2020).

195. Zlotnick, A. *et al.* In vitro screening for molecules that affect virus capsid assembly (and other protein association reactions). *Nat. Protoc.* **2**, 490–498 (2007).

196. Li, C., Wang, J. C.-Y., Taylor, M. W. & Zlotnick, A. In vitro assembly of an empty picornavirus capsid follows a dodecahedral path. *J. Virol.* **86**, 13062–13069 (2012).

197. Hanslip, S. J., Zaccai, N. R., Middelberg, A. P. J. & Falconer, R. J. Intrinsic fluorescence as an analytical probe of virus-like particle assembly and maturation. *Biochem. Biophys. Res. Commun.* **375**, 351–355 (2008).

198. Wu, P. & Brand, L. Resonance energy transfer: methods and applications. *Anal. Biochem.* **218**, 1–13 (1994).

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

200. Ansari, M. Z., Alom, S. E. & Swaminathan, R. Ordered structure induced in human c-Myc PEST region upon forming a disulphide bonded dimer. *J. Chem. Sci.* **133**, 26 (2021).

201. Van Der Meer, B. W., Coker, G. & Chen, S.-Y. S. *Resonance energy transfer: theory and data*. (VCH publishers, 1994).

202. Moens, P. D. J., Helms, M. K. & Jameson, D. M. Detection of tryptophan to tryptophan energy transfer in proteins. *Protein J.* **23**, 79–83 (2004).

203. Stray, S. J., Johnson, J. M., Kopek, B. G. & Zlotnick, A. An in vitro fluorescence screen to identify antivirals that disrupt hepatitis B virus capsid assembly. *Nat. Biotechnol.* **24**, 358–362 (2006).