

---

# LOW ENERGY OPTICAL EXCITATIONS AS AN INDICATOR OF STRUCTURAL CHANGES INITIATED AT THE TERMINI OF AMYLOID PROTEINS

---

A PREPRINT

KwangHyok Jong<sup>‡,§,‡</sup>, Yavar T. Azar<sup>¶</sup>, Luca Grisanti<sup>§</sup>, Amberley D. Stephens<sup>‡</sup>, Saul T.E. Jones<sup>‡</sup>, Dan Credgington<sup>‡</sup>, Gabriele S. Kaminski Schierle<sup>\*‡</sup>, and Ali Hassanali<sup>\*‡</sup>

<sup>‡</sup>International Centre for Theoretical Physics, Strada Costiera 11, Trieste 34151, Italy

<sup>§</sup>SISSA-Scuola Internazionale Superiore di Studi Avanzati, via Bonomea 265, Trieste 34136, Italy

<sup>¶</sup>Physics and accelerators school, NSTRI, AEOI, P. O. Box 14395-836, Tehran, Iran

<sup>‡</sup>Department of Chemical Engineering and Biotechnology, University of Cambridge, Pembroke Street, Cambridge CB2 3RA, United Kingdom

<sup>‡</sup>Department of Physics, Kim Il Sung University, RyongNam Dong, TaeSong District, Pyongyang, D.P.R., Korea

\*E-mail: ahassana@ictp.it, gsk20@cam.ac.uk

October 16, 2018

## 1 Introduction

Amyloid aggregation is a process by which proteins assemble into fibrils characterized by a high content of  $\beta$ -sheets[1, 2, 3]. The formation of fibrils is associated with neurodegenerative diseases such as Alzheimers and Parkinsons[4, 5]. In the past decades, there has been a lot of experimental evidence showing that small oligomers and protofibrils in the early stages of fibril aggregation, are in fact, more toxic[5, 6, 7, 8]. An understanding of the physical and chemical processes associated with the initial stages of amyloid aggregation may help with coming up with medical therapies that can intervene at earlier stages of the disease development.

The stability of amyloid fibrils has typically been rationalized by the presence of a dense network of hydrogen bonds forming  $\beta$ -sheet structures[2, 9, 10]. Recent experimental using solid-state NMR spectroscopy however, show that 3D structure of  $A\beta_{1-42}$  fibrils is also characterized by a hydrophobic core made up of maximally buried hydrophobic side chains[11]. Moreover, it has also been shown that the toxicity of the oligomers increases with increasing hydrophobicity[6, 7, 8]. The interplay between the hydrophobic interactions between the side chains and the hydrophilic interactions through the dense hydrogen bond network and its role in amyloid aggregation still remains an open question.

One of the obvious experimental challenges in characterizing the mechanisms associated with aggregation like those discussed earlier, is being able to monitor the aggregation process. In this regard, there has been an increasing effort to understand the optical properties of amyloid proteins. Recent experimental studies have shown that amyloid fibrils develop an intrinsic fluorescence during aggregation[12, 13, 14, 15]. It has also been observed that these protein aggregates can absorb low energy photons in the energy range of black 240-500 nm [16] Remarkably, these features do not require the presence of aromatic amino acids.

These anomalous optical properties are not exclusive to amyloid aggregates[17, 18, 19, 20]. Fluorescence experimental studies for a series of non-aromatic biogenic and synthetic peptides based on alanine, valine and isoleucine also shos that the intrinsic fluorescence in the aggregated state or condensed phase is associated with the abundant existence of hydrogen bonding between amide groups. [18] Recently, Prasad et al.[19] showed that a monomeric protein devoid of aromatic residues features significant absorption between 250-300nm and a long tail in the absorption up to about 800nm. Using electronic structure calculations, they show that the charged amino ( $\text{NH}_3^+$ ) and carboxylate ( $\text{COO}^-$ ) of the spatially proximal lysine and glutamic acid side chains, act as electronic charge acceptors or donors for photoinduced electron transfer either from or to the polypeptide backbone or to each other. They also demonstrate that charge-transfer (CT) excitations involving these charged groups and the peptide backbone appear to be the main source for the optical

activity of these proteins in the range between 220-380 nm. In specific cases, the increase of the intrinsic fluorescence during the aggregation of insulin and lysozyme suggest that this phenomena are at least partially caused by the chemical process like oxidation [21].

It has also been suggested that this intrinsic fluorescence might be utilized as a label-free diagnostic tool to probe the structural and dynamical transition of amyloid-like aggregates. For example, Ansari et al.[22] showed that the absorption intensities of the PEST fragment of human c-Myc and its mutant at wavelengths of 250-800 nm are dependent on the 3D proximity of the charged functional groups across the protein. Moreover, the significant changes in the spectra by changing pH in the range of 3-11 as well as by the application of different temperatures and salts show a strong correlation between their secondary structure and fluorescence. The increase of the absorbance spectra with time in Hen Egg-White Lysozyme at pH of 2 directly correlated with the growth of aggregates, as confirmed by the increasing thioflavin T fluorescence. Meanwhile, the sensitivity of absorption spectra by the charge transfer transition to the proximity of the different charged groups is also very supportive to the feasibility to employ the non-aromatic residues as a new luminescent biomolecules[23, 19, 20].

In this work, we combine absorption and fluorescence spectroscopy experiments and theoretical modeling to specifically examine the role of termini interactions on the optical properties. Optical absorption and fluorescence is measured for a six-chain amino acid 2Y3J (AIIGLM) which forms a segment of the full amyloid beta 1-40. In order to explore the sensitivity of the optical properties to the termini interactions, the experiments were repeated by acetylating the N-terminus. Although atomic force microscopy experiments indicate the formation of some form of fibrillar or crystal aggregates in both systems, the optical properties are strikingly different - acetylation significantly reduces optical activity between 280-350 nm.

In the last couple of years, several of us have been working to combine both theory and experiments to understand the anomalous optical properties of these systems. In earlier studies, we have shown using density functional theory (DFT) based calculations, that strong hydrogen bonding interactions between the N and C termini of the amyloid proteins, play an important role in tuning their optical properties[14]. Due the complexity of the amyloid fibril morphology, these calculations were performed using standard generalized gradient approximation (GGA) functionals on small model crystal structures. Furthermore, in these studies, the role and importance of structural and conformational heterogeneity on the optical properties was not examined. When used to determine optical properties, GGA functionals can often suffer from inadequately capturing the physics of CT excitations. Here, we use range corrected hybrid functionals to examine the electronic character of the optical excitations. We suggest that the low energy excitations observed in the experiment appear to be modulated by interactions involving the termini groups which in turn, is affected by the specific conformations of the peptide. The low energy absorption of photons involves a mixture of CT excitation from *C-to-N* terminus and *C-terminus* to backbone groups. This region of the optical spectrum may thus provide a way to examine the extent of packing and specific hydrogen bond interactions in peptide aggregates.

The remainder of the paper is organized as follows: Section II devoted to a brief explanation of experimental methods and employed computational methods for the structural preparations and ground (excited) states calculations. In section III, we present our results for both ground and excited states as a function of structural parameters, emphasizing the role of all possible types of intramolecular interaction.

## 2 Methods

### 2.1 Experimental Methods

#### 2.1.1 Sample preparation

Six amino acids from the aggregation prone region 30-35 of the Amyloid beta ( $A\beta$ ) protein, encoding AIIGLM, were purchased from Eurogentec, (Belgium) in an NC form (H-AIIGLM-OH) and an N-acetylated form (Ac - AIIGLM - OH) with 97 % purity, determined by high performance liquid chromatography (HPLC). The peptides were diluted to 600  $\mu$ M in freshly made phosphate buffered saline (PBS) (137 mM NaCl, 10 mM K<sub>2</sub>HPO<sub>4</sub>, 2.7 mM KCl, pH of 7.4 (Sigma, UK)) in 18.2  $\Omega$  dH<sub>2</sub>O (MilliQ, Merck Millipore, Germany). 5 mL of each peptide solution was sonicated at 70% power using a Digital Sonifier<sup>®</sup> SLPe (model 4C15, Branson, USA) for 10 seconds with a 10 second rest, repeated three times. Samples were incubated in a 15 mL falcon tube (Corning, USA) under continuous rotation at maximal speed in a UVP HB-1000 Hybridizer (Fisher Scientific, UK) at 37°C for seven days.

#### 2.1.2 Absorption

3 mL of each 600  $\mu$ M NC and N-acetylated peptide solution was analysed in a 10 mm QX cuvette (Hellma, Germany) using a UV-Vis-NIR Spectrophotometer (model UV-3600 Plus, Shimadzu, Japan). Measurements were taken between

200 - 800 nm using 1 nm steps at a slow scan speed and a 1 nm slit width. The light source change wavelength was set at 393 nm and the grating change wavelength set at 750 nm. Samples were measured three times and the experiments repeated three times, measurements were then averaged and PBS only control was subtracted. Data was then processed for scatter subtraction using equation 1 where  $A$  is absorption detected,  $c = 4 \times 10^8$ . Data was presented as molar absorptivity.

$$blackA'(\lambda) = Ablack(\lambda) - \log\left(\frac{1}{1 - c\lambda^{-4}}\right) \quad (1)$$

### 2.1.3 Excitation and Emission Scans

The same samples measured for absorption were also measured for excitation and emission spectra. Excitation and emission measurements were taken in an F-4500 FL Spectrophotometer (Hitachi High-Technologies Corporation, Japan) in 10 mm QX cuvettes. Scans were taken using a 5 nm or 10 nm excitation slit and a 20 nm emission slit. The scan speed was set to 240 nm/min and the PMT voltage set to 950 V. Excitation scans were taken between 250 - 370 nm, with the emission set at 420 nm and emission scans were taken between 380 nm - 560 nm with the excitation set at 305 nm. Four scans were taken of each sample and the experiments were repeated three times, measurements were then averaged and the PBS only control subtracted.

### 2.1.4 Atomic Force Microscopy (AFM)

AFM was used to determine the morphology of formed NC and N-acetylated peptide species after seven days. 200  $\mu$ L of peptide was deposited on freshly cleaved mica and incubated for 15 minutes. The mica was washed twice with 18.2  $\Omega$  dH<sub>2</sub>O and images were acquired in dH<sub>2</sub>O. Images were acquired in tapping mode on a BioScope Resolve (Bruker, AXS GmbH) using ScanAsyst-Fluid+ probes. 256 lines were acquired at a scan rate of 0.966 Hz per image of 2-5  $\mu$ m field of view, for at least six fields of view. Images were adjusted for contrast and exported from NanoScope Analysis 8.2 software (Bruker).

## 2.2 Computational Methods

In the following section, we describe the computational methods that were performed to help interpret the observations made in the experiments. An important aspect of this work was to be able to use more accurate time-dependent density functional theory (TDDFT) functionals in order to characterize the nature of the optical excitations and the differences observed in the experiments. In order to achieve this, the model systems used for the theory need to be carefully identified which we will now describe.

### 2.2.1 Construction of Model Systems

The experimental systems studied for both NC and N-Acetylated involve very large-scale molecular aggregates that are currently computationally prohibitive to model. Furthermore, the morphology of both the fibril or crystal structures in the case of NC and N-Acetylated respectively, is currently not fully characterized. In the case of NC, there is a crystal structure available[24] which formed the study of a previous work in group using GGA functionals[14, 15]. However, the unit cell is made up of 752 atoms and thus one cannot use more accurate range-corrected hybrid functionals to study optical properties on this system. We thus focus our efforts on investigating the electronic structure and optical properties of a monomer unit of this structure where the role of the conformation of the peptide can be examined. There are two important motivations for this: firstly, as pointed out in the introduction, very recent experiments have shown that the fold of  $A\beta_{1-42}$  is stabilized not only by the well known polar interactions of the beta-sheets, but also the hydrophobic packing of the side chains[11]. In addition, there have also been some studies showing that monomeric proteins lacking aromatic amino acids can also exhibit surprising optical properties[19].

We computed the optical spectra using TDDFT for configurations of NC that were obtained from an earlier work in our group[9]. We briefly describe some of the details here. We performed microsecond simulations of the NC chain in a box of water using well-tempered metadynamics simulations which allowed for a thorough sampling of the conformations of the chain. These simulations showed that there were at least four important collective variables involved in characterizing the various structures: end-to-end distance ( $d_{ee}$ ), radius of gyration (Rg), backbone contacts (BC) which quantifies polar interactions within the chain and finally the sidechain contacts (SC) which measures the extent of hydrophobic packing involving the sidechain groups of the amino acids (see SI for the details of the definition of BC and SC). The free energy landscape along  $d_{ee}$ , BC and SC is shown in Figure 1. A total of 40 configurations were chosen in order to determine the optical spectra. These different configurations were chosen from different regions

of the free energy surface (FES) characterized in earlier work[.]. To help with the discussion later on, we define different basins (we will refer to A, B, C and D), which are black all shown in Figure 1.

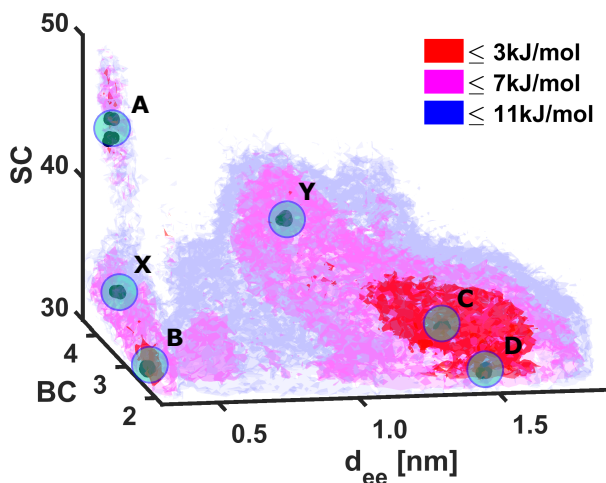


Figure 1: 3D free energy surface along  $d_{ee}$ , BC and SC for NC system.

In order to examine the sensitivity of the optical properties to the acetylation of the N- and C- termini, we performed some calculations on two other model systems. The first one, dubbed ACE, has N-Acetylated black that results into negatively charged C-terminus and is thus likely black to be complexed with a positively charged counterion from the buffer solution. Two extended conformations obtained from basins C of NC were used to acetylate the N-terminus after which a sodium ion was placed near the negatively charged C-terminus. These configurations were then black partially optimized, keeping the backbone heavy atoms frozen. The second model system is inspired from an earlier work in our group[9] where we conducted simulations of the NC system modified so that both the N and C termini were capped with methyl groups. This system is dubbed as MET. As expected the FES of MET along the coordinates illustrated in Figure 1 is qualitatively different (see SI Figure S2) due to the absence of the strong salt-bridge interactions between the termini. While we currently have no experiments on this particular system, it provides an instructive example of the importance of the N-C termini black and its hydrogen bond) interactions on the spectra.

## 2.2.2 Electronic Structure Calculations

All the optical properties are performed using TDDFT[25, 26] implementation in the Gaussian 09 code[27]. In all the ground and excited states calculations, we used a range corrected exchange-correlation functional (CAM-B3LYP)[28] with the 6-311++G\*\*[29, 30] basis set. There have been numerous previous studies showing that CAM-B3LYP performs quite well in characterizing charge transfer excitations as well as Rydberg states[31, 32].

Here, we briefly review the formalism behind computing the optical properties. Excitation energies and oscillator strengths were calculated solving the Casida equations,

$$\sum_{jb} \epsilon_{ia,jb} F_{jb,I} = \omega_I^2 F_{ia,I} \quad (2)$$

where  $\omega_I$  is the excitation energy of  $I$ -th excited state and  $F_{ia,I}$  correspond to the contribution of the transition from an occupied KS orbital ( $\phi_i$ ), to a virtual orbital ( $\phi_a$ )[33], where  $i,j,\dots$  and  $a,b,\dots$  indices used for occupied and virtual orbitals, respectively. The solution to the equation yields the corresponding oscillator strengths for each excitation through the following:

$$f_I = \frac{2}{3} \omega_I |d_I|^2, \quad d_I = \sum_{ia} \sqrt{\frac{2\omega_{ia}}{\omega_I}} F_{ia,I} \vec{d}_{ia} \vec{d}_{ia} = \langle \phi_i | \vec{r} | \phi_a \rangle \quad (3)$$

The transition dipole moment for the  $I$ -th excitation,  $d_I$ , can be expressed as a linear combination of single orbital transitions,  $d_{ia}$ . In order to obtain the spectra, these resulting oscillator strengths,  $f_I$ , are convoluted with Gaussian

functions to obtain the molar absorptivity ( $\epsilon$ ) as a function of frequency ( $\omega$ ) in the following manner:

$$\epsilon(\omega) = 2.174 \times 10^8 \sum_I \frac{f_I}{\Delta_{1/2}} \exp\left[-2.773 \frac{(\omega - \omega_I)^2}{\Delta_{1/2}^2}\right] \quad (4)$$

The full width at half maximum ( $\Delta_{1/2}$ ) is 0.3 eV and  $f_I$  and  $\omega_I$  are the  $I$ th oscillator strength and excitation frequency, respectively.

In order to characterize the physical origin of the excitations, charge density differences,  $\Delta\rho(\mathbf{r})$ , are determined for some specific excitations. The spatial distance between the barycenter of positive ( $\Delta\rho_+(\mathbf{r})$ ) and negative ( $\Delta\rho_-(\mathbf{r})$ ) regions can be used to quantify the charge transfer character of each excitation [34]. In addition, we also studied the dipole moment difference between the ground and excited states: [35]

$$\Delta\mu = \int \mathbf{r}\rho_{exc}d\mathbf{r} - \int \mathbf{r}\rho_{gs}d\mathbf{r} \quad (5)$$

where  $\rho_{gs}$  and  $\rho_{exc}$  are the electronic density of the ground and excited states.

As indicated earlier, we cannot currently perform TDDFT calculations on molecular aggregates of system sizes that are relevant to the experiment. However, the effect of the environment can be included in a mean field manner using polarizable continuum solvent models[36, 37] (PCM). For our calculations, we use a conductor-like polarizable continuum model (C-PCM)[38] implemented in Gaussian. The static and optical dielectric constants for the protein were chosen to be 10 and 2.5 respectively, as used in several previous studies[39, 40].

### 3 Results and Discussion

We begin by first showing the experimental results showing the effect of altering the chemistry of the termini group on the absorption and emission spectra, followed by a detailed analysis of our theoretical findings.

#### 3.1 Experimental Results

The left and right panels of Figure 2 illustrate the AFM images obtained for the NC versus ACE systems. NC appears to form a mixture of both fibrils and other types of aggregates while ACE on the other hand, does not have any signature of fibril formation. The exact morphology of these systems cannot be inferred from the AFM images. Specifically, it is also known that amyloid aggregates are highly polymorphic given the wide variety of structures that have been found using NMR and X-Ray crystallography[41].

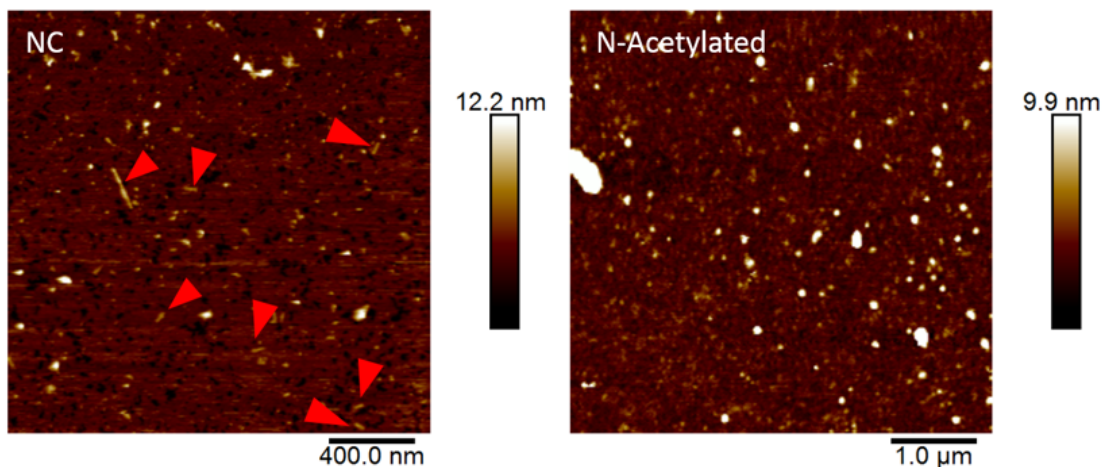


Figure 2: Atomic Force images of amyloid proteins in PeakForce tapping mode. Samples grown at pH7.4: (a)  $A\beta_{30-35}$ , (c) acetylated- $A\beta_{30-35}$ .

Figure 3 shows the absorption spectra for the NC and N-Acetylated systems. Acetylating the N-terminus results in a drastic change in the optical absorption. Although measuring pure absorption spectra is challenging for such samples and scattering contributes to the background, the absorption intensity of the entire low-energy band between 250 and

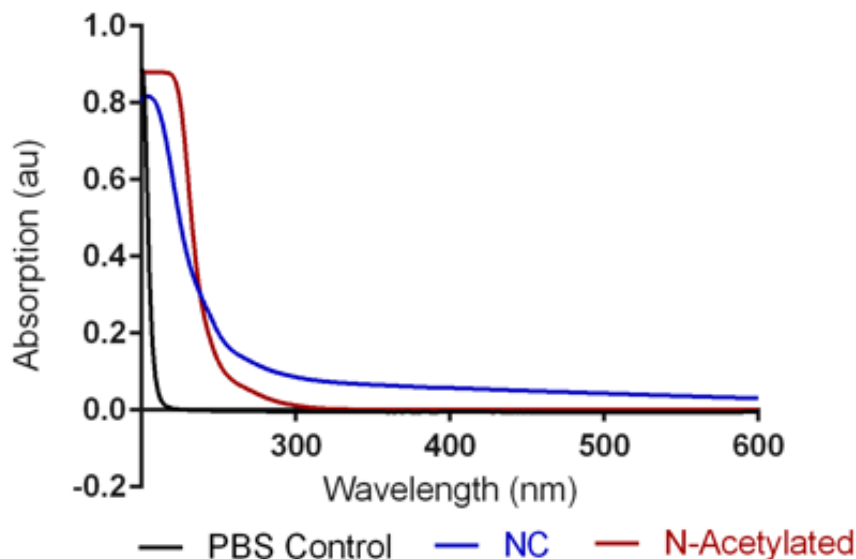


Figure 3: Absorption spectra of  $A\beta_{30-35}$  and acetylated  $A\beta_{30-35}$  compared to pH 7.4 buer background

450nm is significantly reduced in going from NC to ACE and the optical density of the latter overlaps with the noise from the background. Thus the enhanced absorption of the NC system is significant.

Fluorescence excitation and emission spectra were determined for both NC and ACE structures. The excitation was centered at 350 nm and the emission at 420 nm for both systems. Figure 4 shows these spectra for NC and ACE in the left and right panels respectively. Interestingly, for NC both the excitation and emission spectra are very similar to that of the full amyloid[14]. Again, acetylation results in a relative reduction in the low energy component of the fluorescence (near-UV region).

The experimental findings show a striking sensitivity to the presence of having the zwitterionic state of the termini ends and the optical absorption/emission. There are several possible scenarios that could result in this behavior. If the low energy excitation band between 250 and 450nm involves charge reorganization near the N-terminus, it is clear that acetylating could quench this excitation. On the other hand, Figure 2 shows, that the morphology of the amyloid aggregate changes upon N-terminal acetylation. The exact details of how this change is manifested in the aggregate remains unknown. However, the change in optical properties could also be the result of differences in polar and non-polar interactions in the aggregate which ultimately change the environmental conditions and nature of the excitations. black The difference in absorption versus fluorescence excitation profile are also a clear indicator of heterogeneous optical behaviours. In the next section, we will explore using our simulations, the molecular origins of these trends.

### 3.2 Theoretical Results

We begin with our analysis of the electronic structure from the ground-state DFT calculations followed by the optical properties obtained with TDDFT. In particular, we will examine how the the optical properties are affected by the peptide conformation and specific environmental interactions.

#### 3.2.1 Ground State Electronic Structure

In order to get a qualitative sense of the role of conformation and environmental effects on the electronic structure, we begin by showing in Figure 5 the KS energy levels for the occupied and virtual orbitals for all the forty configurations sampled from the FES. The KS energy diagrams in panel (a) of Figure 5 shows a significant difference in the HOMO-LUMO gaps between the *compact* states of A and B and the *extended* states of C and D in the NC system. While the values of HOMO-LUMO gaps for the states A and B are larger than 4 eV, those of states C and D are less than 2 eV. Panel (b) of Figure 5 shows the KS energy diagrams when the protein environment is included using the PCM model.

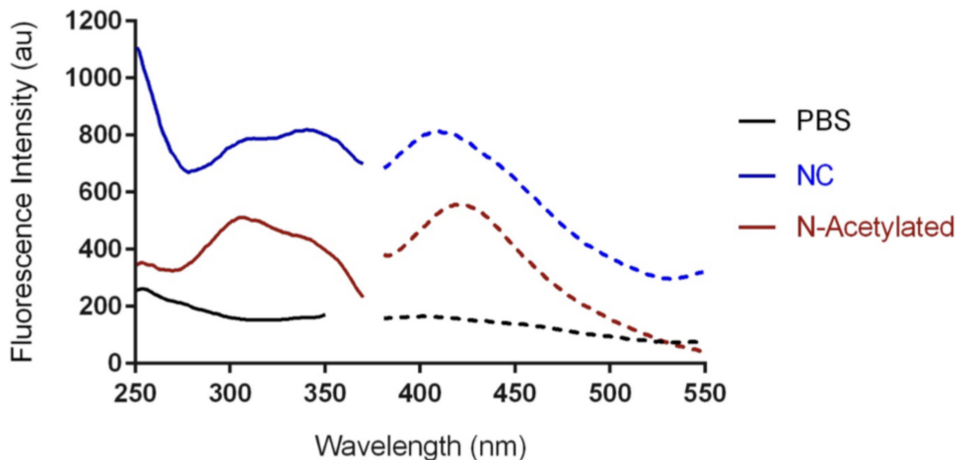


Figure 4: Emission and excitation spectra of the intrinsic fluorescence of amyloid protein solutions: (a)  $A\beta_{30-35}$ , (b) acetylated  $A\beta_{30-35}$ . For the excitation spectrum, the emission was collected at 350 nm, and for the emission spectra the excitation was set at 290 nm.

The HOMO-LUMO gaps for the compact states change by a very small amount of  $\approx 0.1$  eV, for the extended states the gap increases by  $\approx 4$  eV.

Iida et al.[42] have discussed a systematic framework for understanding the orbital energy shifts that arises due to the effect of solvation. The orbital energy shift of  $i$ th level,  $\Delta\epsilon_i$ , is proportional to the dot product of the change in the dipole moment associated with  $i$ th molecular orbital ( $\Delta\vec{\mu}_i$ ) and total dipole moment of the molecule ( $\vec{\mu}_0$ ), i.e.,  $\Delta\epsilon_i \propto -\Delta\vec{\mu}_i \cdot \vec{\mu}_0$ . Figure 6 shows the spatial distributions and dipole moments for HOMO and LUMO as well as the total dipole moment of the representative compact and extended configurations of NC system. The dipole moments of HOMO and LUMO and the total dipole moment are shown in magenta and gray colors, respectively. As seen in the left and right panels of Figure 6, the magnitudes of the dipole moments of the HOMO and LUMO and the total dipole moment of the extended configuration are larger as expected, than ones of the compact configurations. Thus the effect of the solvent environment on the HOMO-LUMO gap is quite sensitive to the specific details of the conformation of the peptide.

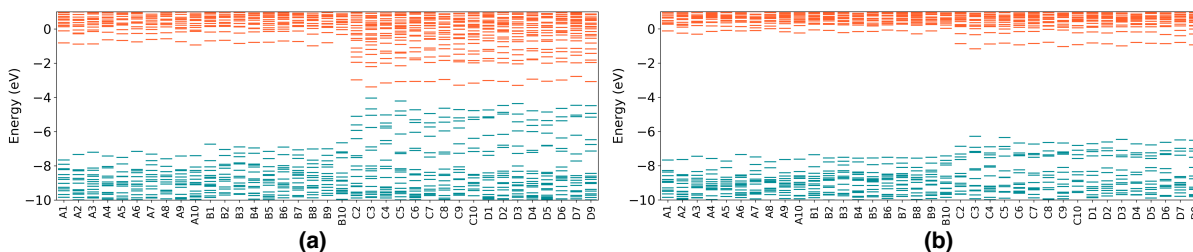


Figure 5: KS energy diagram for the configurations of the states A, B, C and D of NC system (a) in gas phase, (b) in the protein black effective environment. Cyan and orange bars represent HOMO(s) and LUMO(s) respectively.

### 3.2.2 Absorption Spectra

We move on now to examining the effect of the protein conformation on the absorption spectra. Figure 1 shows the free energy surface of the monomer chain from which a series of configurations were sampled to compute the spectra. The left and right panels of Figure 7, show the averaged absorption spectra over the different configurations for calculations performed in the gas phase and using the PCM model, respectively. In order to assess the role of protein conformation on the spectra, we separate the contributions coming from the compact (A and B) and extended states (C and D).

The averaged spectra in the gas phase for all NC states in panel (a) show tails beyond 250 nm and especially the extended states C and D have longer tails beyond 300nm. As we can see in the panel (b), in the protein black effective environment there still remain tails in the spectra at wavelengths larger than 250 nm for the states C and D. However,

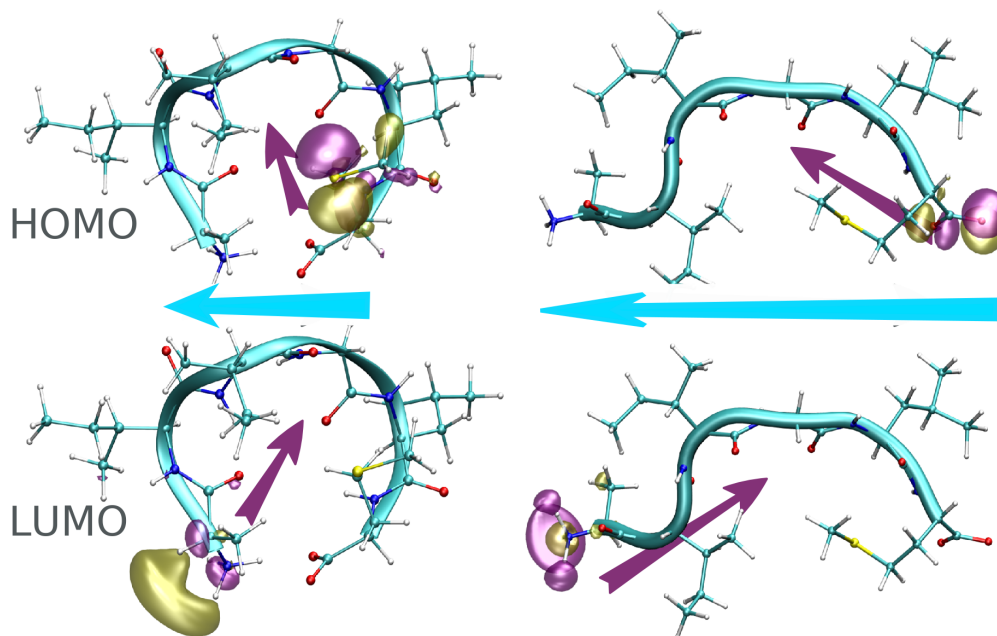


Figure 6: HOMO and LUMO isodensity plots of NC compact configuration (left) and the NC extended configuration (right). The cyan arrows represent molecule dipole moments and magenta arrows represent the dipole moments associated with HOMO and LUMO.

the tails beyond 250 nm vanish for states A and B. Both in the gas phase and PCM environment, the averaged spectra of states C and D exhibit relatively larger values of the molar absorptivity than the ones of states A and B. Interestingly, among the different states of NC system in the PCM model, only state C displays clearly a black shoulder peak at a large wavelength of about 260 nm which is very close to the experimental absorption peak, suggesting the dominance of more extended like conformations in the fibril. This feature is also consistent with our previous studies where the optical spectra were determined for the fully periodic 2Y3J crystal, albeit with a lower level of theory[14].

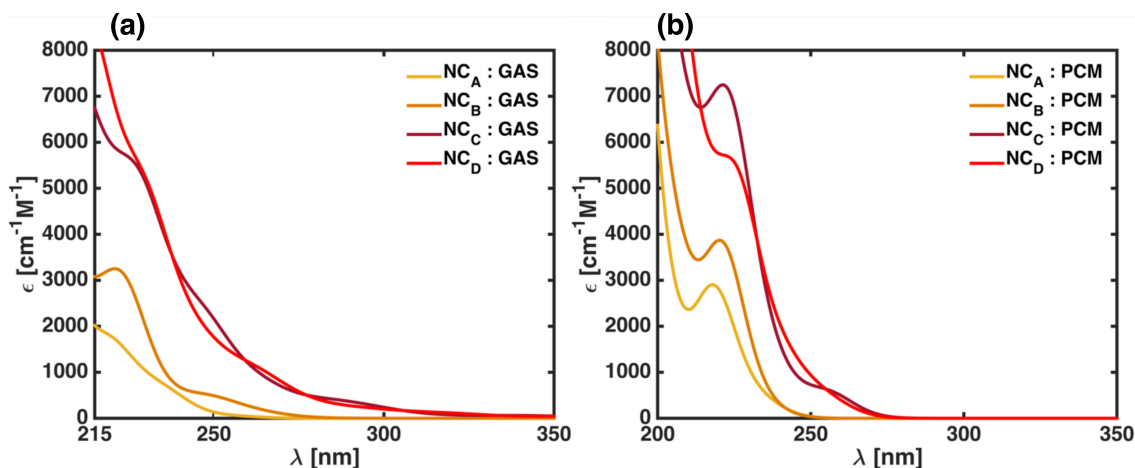


Figure 7: Averaged absorption spectra on different states of NC system in gas phase (a) and NC system in the protein black effective environment (b).

Since our absorption spectra are constructed from configurations sampled from a molecular dynamics simulation, we are also able to investigate the coupling between the optical properties and the conformation of the peptide. Since the conformational fluctuation of the protein can be generally mapped to the changes in the extent of the polar and non-polar interactions of the backbones and side-chains, we estimated the backbone and side-chain contacts (BC and



SC) of the peptide. Note that the BC refers to all the intra hydrogen bonds including the interactions between the N and C termini.

For all the configurations of NC, we explored the correlation between the positions of the first peak in the black average absorption spectra,  $E_1(\lambda)$ , with BC and SC. In Figure 8, we show the 2D scatter plots between  $E_1(\lambda)$  and the structural parameters in the gas phase. The size of the scatter points is representative of the relative oscillator strength of the first peak in the absorption spectra. Both panels a) and b) in Figure 8 show a sizable correlation between structural parameters and the positions of the first peak. Lower BC and SC which correspond to more open or extended conformations are characterized by lower excitation energies.

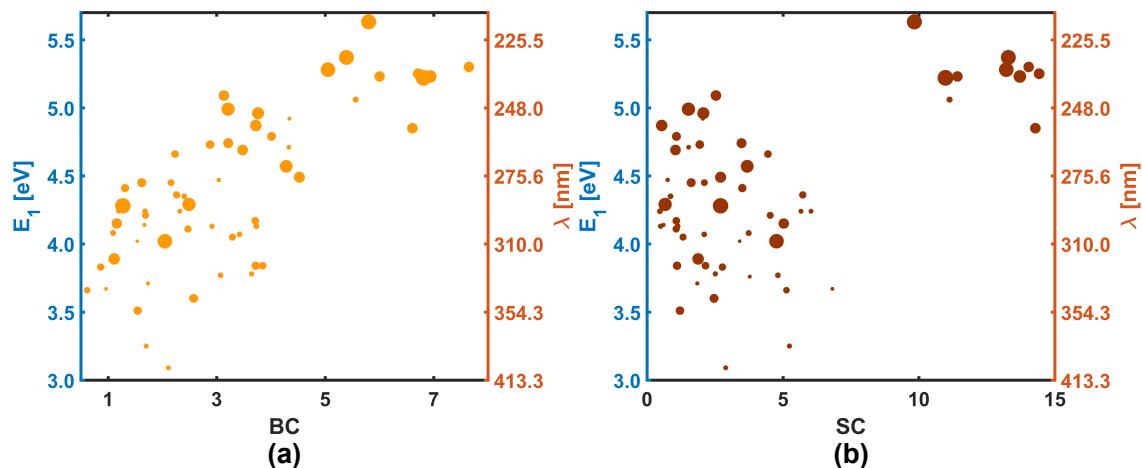


Figure 8: Scatter plot between (a) BC and  $E_1(\lambda)$  and (b) SC and  $E_1(\lambda)$  for NC system. The size of scatter points are representative of the oscillator strength of the corresponding first peak. The values of Pearson's correlation coefficient for two panels in Figure 8 are 0.74 and 0.54 respectively.

From absorption spectra in Figure 7 we have already noticed that an inclusion of the protein environment with the PCM changes the optical nature of the excitations quite significantly. In the panels of Figure 9, the coupling between the BC and SC parameters and the optical properties in the presence of the PCM is illustrated. Interestingly, the correlation with BC is still quite pronounced (with a correlation coefficient of 0.50) and confirms that the low energy excitations in the spectra are associated with conformations that have the least intra-protein hydrogen bond interactions. On the other hand, for the SC coordinate, correlation is weakened in the PCM environment.

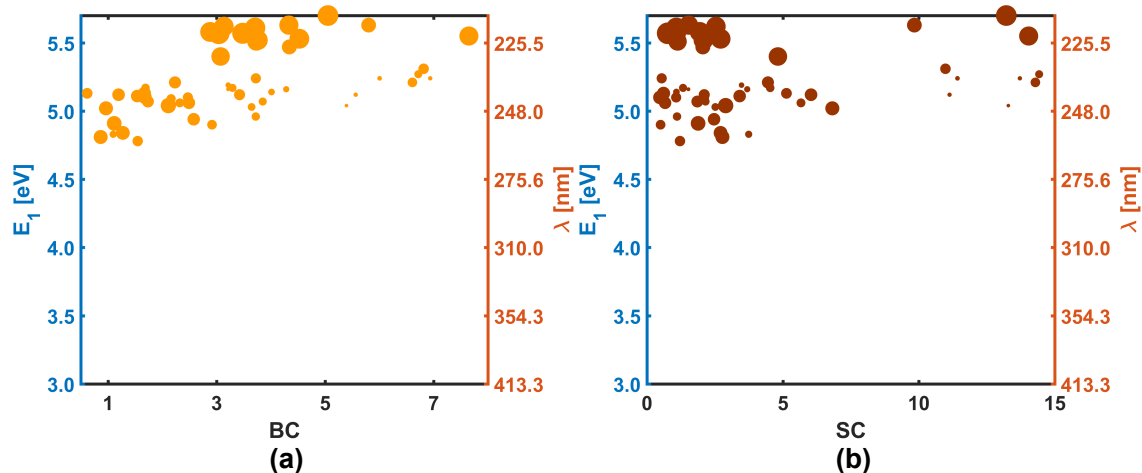


Figure 9: Scatter plot between (a) BC and  $E_1(\lambda)$  and (b) SC and  $E_1(\lambda)$  for NC system in PCM model. The size of scatter points are representative of the oscillator strength of the corresponding first peak.

### 3.2.3 Characterization of Excited States

For polypeptides and proteins without any aromatic residues, it is well known that a sequence of peptide bonds exhibits strong absorption at around 190 nm and weak absorption around 220 nm[43]. Furthermore, optical activity at wavelengths beyond 250 nm is rarely expected. Absorption between 250 and 290nm is typically associated with the presence of aromatic amino acids such as tryptophan. The origins of low energy excitations we observe may be more complicated than expected. Earlier studies in our group have shown that amyloid fibrils devoid of any conjugated systems can absorb and emit low energy photons in the energy range beyond 350nm. Prasad and co-workers have also recently shown[19] the monomeric proteins lacking aromatic residues show absorption at wavelengths between 250-300 nm. They ascribe this to charge transfer excitations involving charged amino acids.

As discussed earlier, the configurations of NC show a broad absorption spectra with a tail stretching between 250 to 300nm. In order to understand the nature of these lowest energy excitations, we chose some representative configurations for both compact (A3) and extended states (C3 and C7). The integer values correspond to different configurations that were sampled from the FES. Note that C3 makes a dominant contribution to the first peak in the averaged absorption spectra of state C beyond 250 nm using PCM (see Figure S3 in the Supporting Information).

We begin by summarizing in Table 1 the characteristics of the three low-lying excited states for the chosen three configurations. Table 1 presents several important points regarding the nature of the low-energy transitions, including the magnitude of the dipole moment difference between the ground and excited states ( $\Delta\mu$ ), and the distance between electron center ( $\rho_e^-$ ) and hole center ( $\rho_h^+$ ) ( $\Delta r$ ) respectively (see Methods Section earlier for more details).

As we can see in Table 1, the lowest excited states S1 of A3 in the gas phase have absorption frequencies of less than 250 nm. In both cases, the magnitude of the dipole moment difference,  $|\Delta\mu|$ , is larger than 2.5 (eÅ) so that this excitation can be considered as a short range CT excitation. Concurrently, the distance between the electron and hole centers is  $\approx 2$  Å. Note that a carbon-carbon covalent bond length is 1.2 Å and hence this excitation really involves a very local shifting of charge. A visualization of the difference between ground and excited state densities ( $\Delta\rho = \rho_{exc} - \rho_{gs}$ ) as seen in Figure 10 (A3) confirms that the excitation has a CT character involving the N and C termini (including the methionine side chain) in the gas phase. Note that the red and blue colors represent depletion and increment of charge density, respectively.

Upon the inclusion of the PCM dielectric environment, we observe distinct changes in the excitation character. For the compact conformation,  $|\Delta\mu|$  is now less than 1.5 (eÅ) implying the the environment quenches the spatial extent of the excitation. Unlike in the gas phase where the excitation involves a CT from the C to N terminus, with the PCM, the excitation is now a charge reorganization involving the methionine side chain as well as the peptide backbone in agreement to some of our previous findings [44].

Table 1: The characterization of the excited states for the configurations of A3, C3 and C7 in the gas phase and the protein environment. The first column is the type of the environment, the second column is the configuration name, third column is the position of the first peak, the fourth and fifth columns are the index of the excited states and their contribution percentage on the first peak, the sixth column is the oscillator strength, seventh column is the magnitude of the dipole moment difference between the excited state and the ground state, and the last column is the distance between the electron center and hole center through the excitation block (see Methods Section for more details) .

Env.	Conf.	$\lambda$ (nm)	state	Perc.(%)	$f_I$	$ \Delta\mu $ (eÅ)	$\Delta r$ (Å)
7*GAS	4*A3	4*237.1	S1	52.4	0.008	2.64	1.99
			S2	19.6	0.003	1.62	1.42
			S4	11.4	0.002	0.30	0.39
			S6	11.3	0.002	0.42	0.52
	2*C3	2*324.6	S47	67.0	0.002	7.36	4.88
			S43	16.3	0.001	9.34	5.82
	C7	345.4	S14	90.3	0.004	5.69	3.62
6*PCM	4*A3	4*233.1	S1	37.9	0.007	1.45	1.40
			S4	23.1	0.004	0.46	0.58
			S3	19.1	0.003	0.30	0.40
			S2	17.8	0.003	0.42	0.53
	C3	257.8	S2	99.9	0.026	1.05	0.99
	C7	251.0	S1	94.2	0.017	3.06	2.49

Similar features involving the CT character of the lowest energy excitation holds for the extended structures of C3 and C7. In the gas phase, all of the excitations lie beyond 270 nm especially that of C7 which lies above 340 nm. The

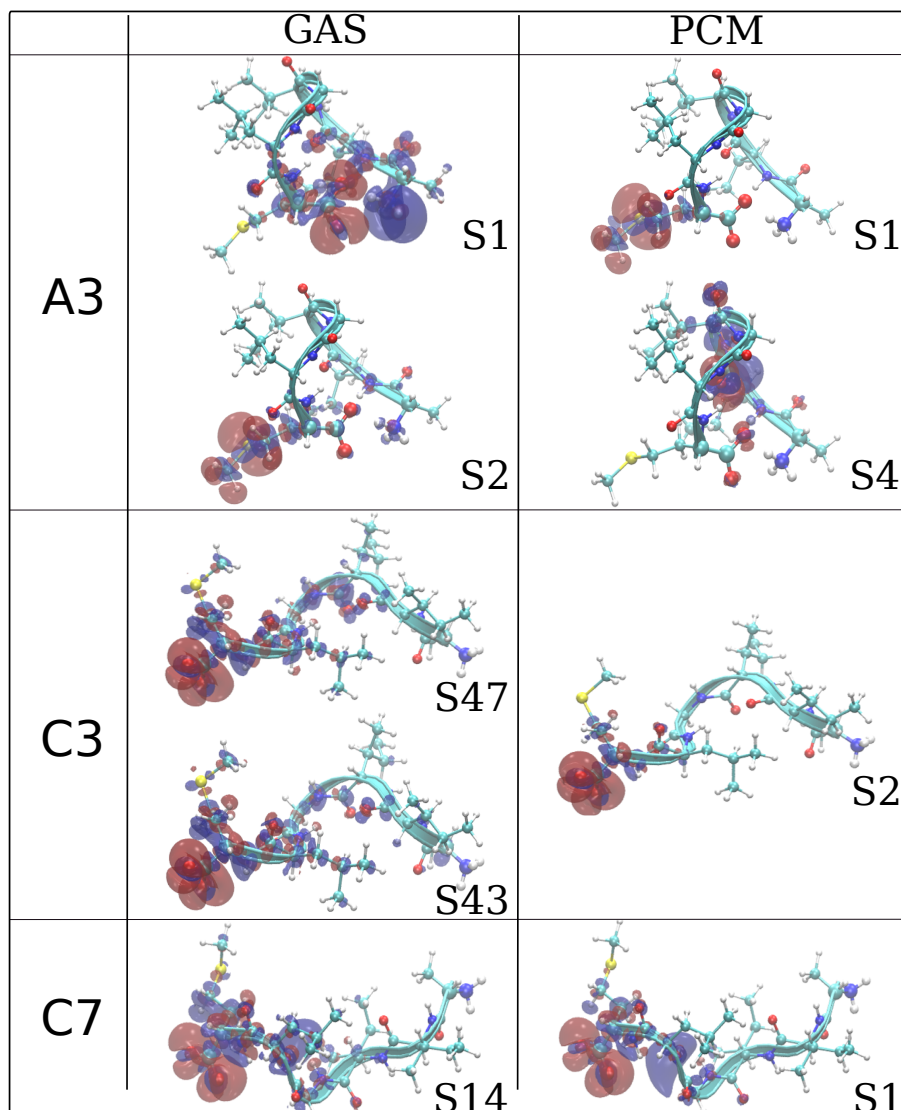


Figure 10: Iso-surface plot of charge difference between ground and excited states. The red and blue colors represent positive (electron) and negative (hole) values of the iso-surface which are  $\pm 0.002$  a.u.

values of  $|\Delta\mu|$  which are larger than  $5.69$  (eÅ) for the low-lying states, show that these excitations have a CT character from the C-terminus to the backbone. Again, this is reflected in the isosurface plots shown in Figure 10. The possibility of charge transfer from a charged terminus to the backbone at wavelengths beyond 250 nm has also been reported previously in small dipeptides[19]. The inclusion of the PCM environment has a significant impact on the character of the excitations. For C3 and C7, we observe that the values of  $|\Delta\mu|$  decrease by 6 (eÅ) and subsequently involve a local CT from the C terminus to the backbone.

### 3.2.4 Role of NC-Termini and Inter Hydrogen Bonding Interactions

The experiments show that upon acetylation of the N terminus, there is a significant decrease in the optical density at low energies. The AFM images also indicate that the aggregated state upon acetylation appears to be quite different from the zwitterionic form, although the exact details of the molecular structure remain unknown at this point. In order to assess the importance of the termini groups on the optical properties, we examined the spectra for two different systems, black defined above. Briefly, we extracted several configuration and performed average from our long MD simulations on MET system (where both termini ends acetylated) and we run constrained optimization on two

selected extended configurations from NC to obtain ACE (N terminus acetylated with a positive counterion ( $\text{Na}^+$ ) near the C-terminus to maintain charge neutrality).

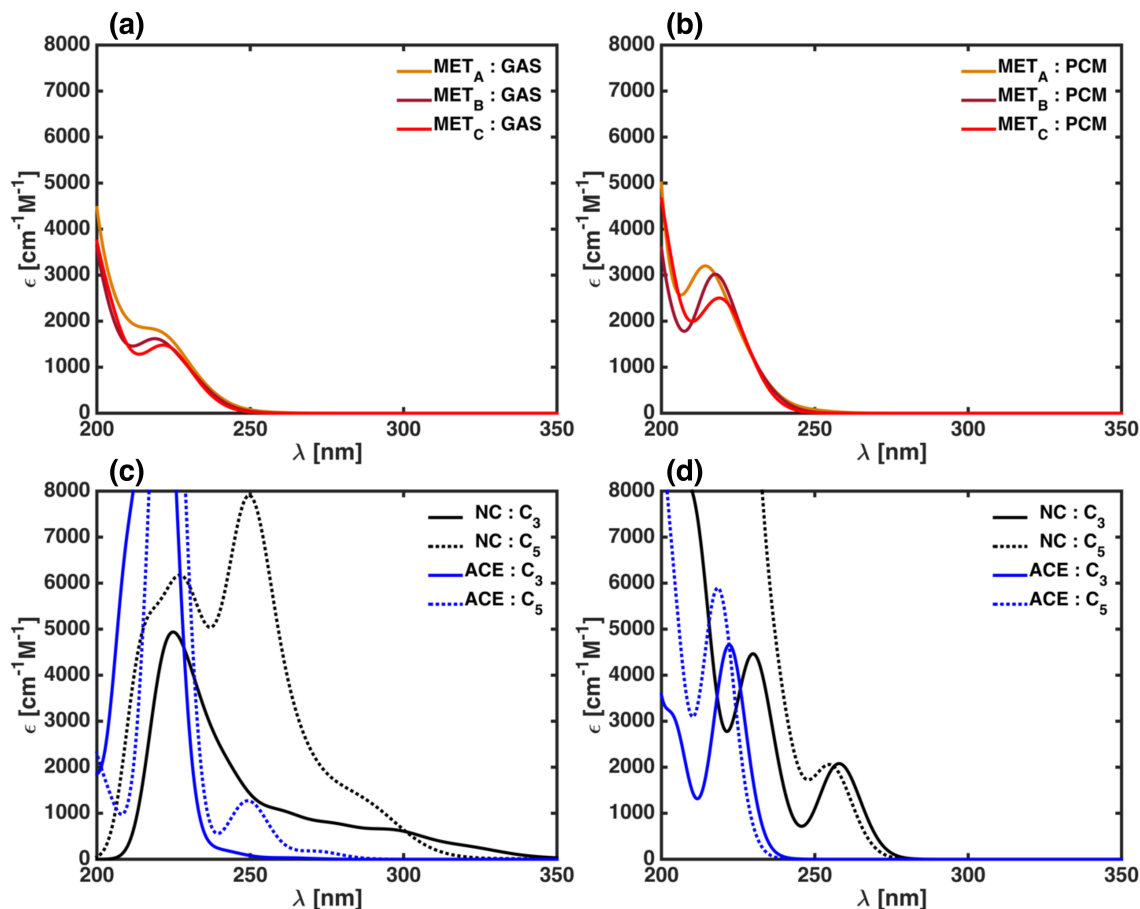


Figure 11: Averaged absorption spectra on different states of MET in the gas phase (a), MET in the protein environment (b), ACE with a sodium counterion in the gas phase (c) and ACE with a sodium counterion in the protein environment (d).

The top left and right panels of Figure 11 show the spectra for MET in the gas phase and in PCM. The results are quite striking when compared to that obtained in Figure 7 for NC: in the case of MET, the region beyond 250nm is dark and furthermore, the spectra in the low-lying energy region are independent of the environment of the peptide. The double-methylation black acetylation is of course a drastic perturbation that changes both the N and C termini chemistry and does not tell us about the individual contributions of the N and C termini to the excitation.

The bottom two panels of Figure 11 show the spectra obtained for two configurations black of ACE (where the N terminus was acetylated), as black measured in the experiment. black For comparison, corresponding NC unperturbed configuration are also shown in the same figure. It is clear that acetylation of the N terminus results in an overall blue shift of the spectra, in both gas phase and PCM model, and that this feature is much more pronounced when the PCM model is applied. The iso-surfaces of charge density difference between the ground states and excited states for the black ACE configurations show the dominant contribution of the local excitations near the sulfur atom and the peptide bonds in the PCM model (see the Fig. SI-10 for the details).

All the analysis up to this point examines the effect of altering the termini of a single chain. We are obviously aware that what we infer about the nature of the optical properties from this system, even with a continuum dielectric medium, is likely to be more complicated black and heterogeneous in the realistic fibrilsblack/aggregate or crystal structures. In order to assess the sensitivity of our interpretations to the chosen model systems black over possible higher hierarchy structures, three dimer configurations were generated using the extended conformation of 2Y3J: a) is a head to tail in one direction where the two beta sheets only interact via the NC termini hydrogen bonds, b) involves two beta sheets in

an anti-parallel configuration and c) involves two beta sheets oriented in a parallel configuration. These three systems are illustrated in Figure 12

Due to the size of these systems, we could not compute the full absorption spectra and instead, we examined only the first few bright excited states. The excitation energies of these states and the corresponding oscillator strengths are shown in Figure 12. In all cases, the protein chains were surrounded by a PCM model using a dielectric constant of 10 to once again, account for the protein environment. Panel (a) and (b) illustrate the head-to-tail and the anti-parallel beta sheet configurations respectively, where, we observe a charge-transfer excitation from the C terminus to the N terminus. In both cases, this excitation involves charge transfer from the C terminus of one chain to the N terminus of another. Interestingly, for the parallel beta-sheet configuration, the excitation is mostly dominated by a short-range charge transfer close to the C terminus. This also happens to be the brightest excitation among those examined in Figure 12.

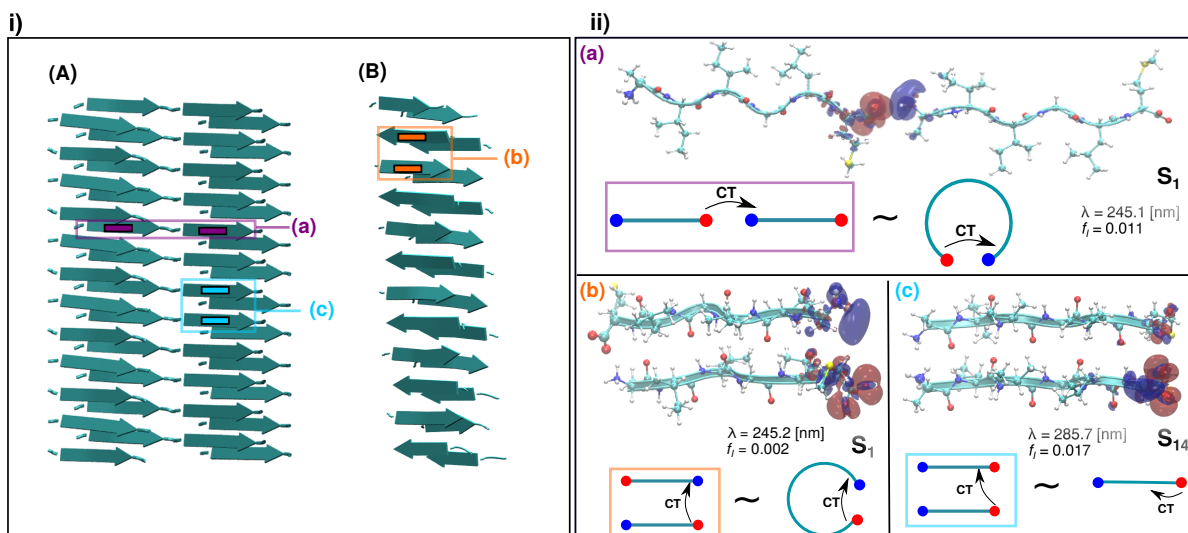


Figure 12: i) Parallel (A) and Anti-parallel (B) protofilament structures ii) Iso-surface plot of charge difference between ground and first excited states for the different dimer configurations : Linear (a), Anti-parallel (b) and Parallel (c). The red and blue colors represent positive (electron) and negative (hole) values of the iso-surface which are  $\pm 0.002$  a.u.. The values of the wavelength( $\lambda$ ) and the oscillator strength ( $f_I$ ) of the lowest energy excitation for the different dimer configurations are shown.

The findings reported in Figure 12 are extremely interesting since it provides clues into linking the effect of conformational changes in the monomer on its optical properties, and the nature of the possible excitations that would occur in a fibril environment. Overall, the effect of changes in conformation in the single chain serves as a good proxy for the more complex interactions that would occur in the fibril, black as reported on right side of Fig. 12 . In a more realistic environment, we can anticipate charge transfer excitations that occur within the same chain and across different chains, the details of which will be sensitively tuned by the conformation of individual chains and how they pack relative to each other.

## 4 Discussion and Conclusions

The question of the origin of low energy excitations and subsequent intrinsic fluorescence in biological systems has been the subject of numerous recent experiments and theoretical studies[14, 15, 19]. In this work, we have tried to focus our efforts on elucidating the role of intra-protein hydrogen bond contacts on optical black properties between 250-300nm beyond the presence of aromatic residues for a some simple model systems.

The fluorescence experiments show a strong sensitivity to the chemical details of the termini. Specifically, acetylation of the N terminus makes the low energy optical activity above 250nm essentially dark. The AFM experiments indicate that both the original zwitterionic chain as well as the N-acetylated chain, appear to form some type of molecular aggregate although the exact details of the structural architecture remain unknown. Using time-dependent density functional theory calculations, we have examined the electronic and optical properties of some model systems of the aggregates both in the gas phase as well as immersed in a continuum dielectric solvent.

Our calculations show that the chemical nature of the termini groups has quite a significant effect on the low energy excitations. In the zwitterionic system, these excitations have a charge transfer character that often involves the termini groups. The nature of these excitations are very sensitive to the environment as well as the conformation of the protein. Our results suggest that the extended conformations are most likely to be the dominant structures in the aggregates for the zwitterionic system although the broad and long nature of the tail in the excitation energies may also be an indication of the heterogeneity in the packing originating from different conformations due to different extents of backbone and side-chain packing. This is somewhat anticipated given recent experimental results of amyloid beta 1-42 showing that an important component of the amyloid structure is the hydrophobic core[].

As in our previous theoretical studies, a quantitative and realistic modeling of the fate of amyloid fibrils upon photoexcitation cannot be currently modeled. We therefore have to rely on changes in optical absorption to track only infer processes that will happen on the excited state. In earlier work[] we examined how specific bond vibrations such as proton transfer and carbonyl stretches were coupled to the optical absorption. Indeed, if these modes lead to conical intersections, their hindrance in the amyloid environment, may provide a mechanism to facilitate fluorescence. On the other side, a non-zero optical activity in absorption is linked to having a transition dipole moment which in the case of the model system studied here, is strongly modulated by the conformation of protein and its subsequent packing. The differences between experimental absorption and fluorescence excitation spectra confirms the presence of conformations with distinct optical properties, featuring a weak but general broad absorption profile in the near-UV, but only a fraction of them exhibiting emissive properties in the same region and producing a characteristic spectral profile in the experimental fluorescence excitation spectrum (see Figures 3 and 4).

Recent studies by Prasad and co-workers combining theory and experiments have suggested that the low energy excitation regime between 240-500 nm maybe a way to interrogate the secondary structure of a protein. Our results complement this study and reinforce the notion that there is some very rich optical activity in the far UV/visible range that maybe used as an indirect probe into the conformational signature of the protein. In addition, there is the possibility that this region of the spectrum can be used to study molecular aggregates and in the interactions between them in a rather unique way.

## References

- [1] Fabrizio Chiti and Christopher M. Dobson. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry*, 75(1):333–366, 2006. PMID: 16756495.
- [2] Jason Greenwald and Roland Riek. Biology of amyloid: Structure, function, and regulation. *Structure*, 18(10):1244–1260, Oct 2010.
- [3] Alex Morriss-Andrews and Joan-Emma Shea. Computational studies of protein aggregation: Methods and applications. *Annual Review of Physical Chemistry*, 66(1):643–666, 2015. PMID: 25648485.
- [4] Fabrizio Chiti and Christopher M. Dobson. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry*, 75(1):333–366, 2006.
- [5] Monica Bucciantini, Elisa Giannoni, Fabrizio Chiti, Fabiana Baroni, Lucia Formigli, Jesús Zurdo, Niccolò Taddei, Giampietro Ramponi, Christopher M. Dobson, and Massimo Stefani. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 416:507–511, Apr 2002.
- [6] Jean-Christophe Rochet and Peter T Lansbury. Amyloid fibrillogenesis: themes and variations. *Current Opinion in Structural Biology*, 10(1):60 – 68, 2000.
- [7] Benedetta Mannini, Estefania Mulvihill, Caterina Sgromo, Roberta Cascella, Reza Khodarahmi, Matteo Ramazzotti, Christopher M. Dobson, Cristina Cecchi, and Fabrizio Chiti. Toxicity of protein oligomers is rationalized by a function combining size and surface hydrophobicity. *ACS Chemical Biology*, 9(10):2309–2317, 2014.
- [8] Ali Reza A. Ladiwala, Jeffrey Litt, Ravi S. Kane, Darryl S. Aucoin, Steven O. Smith, Swarnim Ranjan, Judianne Davis, William E. Van Nostrand, and Peter M. Tessier. Conformational differences between two amyloid oligomers of similar size and dissimilar toxicity. *Journal of Biological Chemistry*, 287(29):24765–24773, 2012.
- [9] KwangHyok Jong, Luca Grisanti, and Ali Hassanali. Hydrogen bond networks and hydrophobic effects in the amyloid  $\beta_{30-35}$  chain in water: A molecular dynamics study. *Journal of Chemical Information and Modeling*, 57(7):1548–1562, 2017. PMID: 28603985.

- [10] Tsemekhman Kiril, Goldschmidt Lukasz, Eisenberg David, and Baker David. Cooperative hydrogen bonding in amyloid formation. *Protein Science*, 16(4):761–764.
- [11] Marielle Aulikki Wälti, Francesco Ravotti, Hiromi Arai, Charles G. Glabe, Joseph S. Wall, Anja Böckmann, Peter Güntert, Beat H. Meier, and Roland Riek. Atomic-resolution structure of a disease-relevant a(1–42) amyloid fibril. *Proceedings of the National Academy of Sciences*, 113(34):E4976–E4984, 2016.
- [12] Samir Maji, Jason Amsden, Kenneth Rothschild, Margaret M Condrón, and David Teplow. Conformational dynamics of amyloid -protein assembly probed using intrinsic fluorescence †. 44:13365–76, 11 2005.
- [13] Fiona T. S. Chan, Gabriele S. Kaminski Schierle, Janet R. Kumita, Carlos W. Bertoncini, Christopher M. Dobson, and Clemens F. Kaminski. Protein amyloids develop an intrinsic fluorescence signature during aggregation. *Analyst*, 138:2156–2162, 2013.
- [14] Dorothea Pinotsi, Luca Grisanti, Pierre Mahou, Ralph Gebauer, Clemens F. Kaminski, Ali Hassanali, and Gabriele S. Kaminski Schierle. Proton transfer and structure specific fluorescence in hydrogen bond rich protein structures. *J. Am. Chem. Soc.*, 138(9):3046–3057, 2016.
- [15] Luca Grisanti, Dorothea Pinotsi, Ralph Gebauer, Gabriele S. Kaminski Schierle, and Ali A. Hassanali. A computational study on how structure influences the optical properties in model crystal structures of amyloid fibrils. *Phys. Chem. Chem. Phys.*, 19:4030–4040, 2017.
- [16] Patrik K. Johansson and Patrick Koelsch. Label-free imaging of amyloids using their intrinsic linear and nonlinear optical properties. *Biomed. Opt. Express*, 8(2):743–756, Feb 2017.
- [17] Arpan Bhattacharya, Soumitra Bhowmik, Amit K. Singh, Prashant Kodgire, Apurba K. Das, and Tushar Kanti Mukherjee. Direct evidence of intrinsic blue fluorescence from oligomeric interfaces of human serum albumin. *Langmuir*, 33(40):10606–10615, 2017. PMID: 28930631.
- [18] Ruquan Ye, Yuanyue Liu, Haoke Zhang, Huifang Su, Yilin Zhang, Liguó Xu, Rongrong Hu, Ryan T. K. Kwok, Kam Sing Wong, Jacky W. Y. Lam, William A. Goddard, and Ben Zhong Tang. Non-conventional fluorescent biogenic and synthetic polymers without aromatic rings. *Polym. Chem.*, pages 1722–1727, 2017.
- [19] Saumya Prasad, Imon Mandal, Shubham Singh, Ashim Paul, Bhubaneswar Mandal, Ravindra Venkatramani, and Rajaram Swaminathan. Near uv-visible electronic absorption originating from charged amino acids in a monomeric protein. *Chemical science*, 8(8):5416–5433, 2017.
- [20] Xiaohong Chen, Weijian Luo, Huili Ma, Qian Peng, Wang Zhang Yuan, and Yongming Zhang. Prevalent intrinsic emission from nonaromatic amino acids and poly(amino acids). *Science China Chemistry*, 61(3):351–359, Mar 2018.
- [21] Tatiana N. Tikhonova, Nataliya R. Rovnyagina, Alexander Ya Zhrebker, Nikolai N. Sluchanko, Anna A. Rubekina, Anton S. Orekhov, Eugene N. Nikolaev, Victor V. Fadeev, Vladimir N. Uversky, and Evgeny A. Shirshin. Dissection of the deep-blue autofluorescence changes accompanying amyloid fibrillation. *Archives of Biochemistry and Biophysics*, 651:13 – 20, 2018.
- [22] Mohd. Ziauddin Ansari, Amrendra Kumar, Dileep Ahari, Anurag Priyadarshi, Padmavathi Lolla, Rashna Bhandari, and Rajaram Swaminathan. Protein charge transfer absorption spectra: an intrinsic probe to monitor structural and oligomeric transitions in proteins. *Faraday Discuss.*, 207:91–113, 2018.
- [23] I. Mandal, S. Paul, and R. Venkatramani. Optical backbone-sidechain charge transfer transitions in proteins sensitive to secondary structure and modifications. *Faraday Discuss.*, 207:115–135, 2018.
- [24] Jacques-Philippe Colletier, Arthur Laganowsky, Meytal Landau, Minglei Zhao, Angela B. Soriaga, Lukasz Goldschmidt, David Flot, Duilio Cascio, Michael R. Sawaya, and David Eisenberg. Molecular basis for amyloid-polymorphism. *Proceedings of the National Academy of Sciences*, 108(41):16938–16943, 2011.
- [25] Miguel AL Marques and Eberhard KU Gross. Time-dependent density functional theory. *Annu. Rev. Phys. Chem.*, 55:427–455, 2004.

- [26] Erich Runge and Eberhard KU Gross. Density-functional theory for time-dependent systems. *Physical Review Letters*, 52(12):997, 1984.
- [27] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian09 Revision E.01. Gaussian Inc. Wallingford CT 2009.
- [28] Takeshi Yanai, David P Tew, and Nicholas C Handy. A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp). *Chemical Physics Letters*, 393(1-3):51–57, 2004.
- [29] A. D. McLean and G. S. Chandler. Contracted gaussian basis sets for molecular calculations. i. second row atoms,  $z=11-18$ . *The Journal of Chemical Physics*, 72(10):5639–5648, 1980.
- [30] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople. Self-consistent molecular orbital methods. xx. a basis set for correlated wave functions. *The Journal of Chemical Physics*, 72(1):650–654, 1980.
- [31] Michael JG Peach, Peter Benfield, Trygve Helgaker, and David J Tozer. Excitation energies in density functional theory: An evaluation and a diagnostic test. *The Journal of chemical physics*, 128:044118, 2008.
- [32] Zlatko Brkljača, Momir Mališ, David M Smith, and Ana-Suncana Smith. Calculating cd spectra of flexible peptides: An assessment of td-dft functionals. *Journal of chemical theory and computation*, 10(8):3270–3279, 2014.
- [33] Robert Rürger, Erik van Lenthe, Thomas Heine, and Lucas Visscher. Tight-binding approximations to time-dependent density functional theory—a fast approach for the calculation of electronically excited states. *The Journal of chemical physics*, 144(18):184103, 2016.
- [34] Marika Savarese, Ciro Achille Guido, Eric Bremond, Ilaria Ciofini, and Carlo Adamo. Metrics for molecular electronic excitations: A comparison between orbital-and density-based descriptors. *The Journal of Physical Chemistry A*, 121(40):7543–7549, 2017.
- [35] Yavar T Azar and Mahmoud Payami. Theoretical description of efficiency enhancement in dsscs sensitized by newly synthesized heteroleptic ru complexes. *Physical Chemistry Chemical Physics*, 17(44):29574–29585, 2015.
- [36] Christopher J Cramer and Donald G Truhlar. Implicit solvation models: equilibria, structure, spectra, and dynamics. *Chemical Reviews*, 99(8):2161–2200, 1999.
- [37] Jacopo Tomasi, Benedetta Mennucci, and Roberto Cammi. Quantum mechanical continuum solvation models. *Chemical Reviews-Columbus*, 105(8):2999–3094, 2005.
- [38] Maurizio Cossi and Vincenzo Barone. Time-dependent density functional theory for molecules in liquid solutions. *The Journal of chemical physics*, 115:4708, 2001.
- [39] Lin Li, Chuan Li, Zhe Zhang, and Emil Alexov. On the dielectric “constant” of proteins: Smooth dielectric function for macromolecular modeling and its implementation in delphi. *Journal of Chemical Theory and Computation*, 9(4):2126–2136, 2013. PMID: 23585741.
- [40] Jed W Pitera, Michael Falta, and Wilfred F van Gunsteren. Dielectric properties of proteins from simulation: the effects of solvent, ligands, ph, and temperature. *Biophysical journal*, 80(6):2546–2555, 2001.
- [41] Marielle Aulikki Wälti, Francesco Ravotti, Hiromi Arai, Charles G. Glabe, Joseph S. Wall, Anja Böckmann, Peter Güntert, Beat H. Meier, and Roland Riek. Atomic-resolution structure of a disease-relevant a(1–42) amyloid fibril. *Proceedings of the National Academy of Sciences*, 113(34):E4976–E4984, 2016.



- [42] Kenji Iida, Daisuke Yokogawa, Hirofumi Sato, and Shigeyoshi Sakaki. A systematic understanding of orbital energy shift in polar solvent. *The Journal of chemical physics*, 130(4):044107, 2009.
- [43] HD Hunt and WT Simpson. Spectra of simple amides in the vacuum ultraviolet1. *Journal of the American Chemical Society*, 75(18):4540–4543, 1953.
- [44] Dorothea Pinotsi, Luca Grisanti, Pierre Mahou, Ralph Gebauer, Clemens F. Kaminski, Ali Hassanali, and Gabriele S. Kaminski Schierle. Proton transfer and structure-specific fluorescence in hydrogen bond-rich protein structures. *Journal of the American Chemical Society*, 138(9):3046–3057, 2016.