Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting information for: Low Energy Optical Excitations as an Indicator of Structural Changes Initiated at the Termini of Amyloid Proteins

KwangHyok Jong,[†,‡,‖] Yavar T. Azar,[¶] Luca Grisanti,[‡] Amberley D. Stephens,[§] Saul T.E. Jones,[§] Dan Credgington,[§] Gabriele S. Kaminski Schierle,[*,§] and Ali Hassanali[*,†]

†*Condensed Matter and Statistical Physics, International Centre for Theoretical Physics, Strada Costiera 11, Trieste 34151 Italy*

‡*SISSA-Scuola Internazionale Superiore di Studi Avanzati, via Bonomea 265, Trieste 34136, Italy*

¶*Physics and accelerators school, NSTRI, AEOI, P. O. Box 14395-836, Tehran, Iran*

§*Department of Chemical Engineering and Biotechnology, University of Cambridge, Pembroke Street, Cambridge CB2 3RA, United Kingdom*

‖*Department of Physics, Kim Il Sung University, RyongNam Dong, TaeSong District, Pyongyang, D.P.R., Korea*

E-mail: gsk20@cam.ac.uk; ahassana@ictp.it

SI-1 Collective Variables(CVs)

As we discussed in our previous work[S1], we use the several collective variables to extract the relatively stable structures of Aβ_{30-35} and methyl-capped Aβ_{30-35} in calculating the absorption spectra. As one of the collective variables, we chose the end-to-end distance(d_{ee}), the distance between the partially charged NH_3^+ and COO^- groups because there exists a strong electrostatic interactions between them. Another two collective variables are the number of hydrogen bonds between NH and CO groups of the backbone (BC) and the van-der-Waals interaction interactions between side-chain residues (SC). The BC and SC are calculated by the switching functions defined as

$$\sum_{i\in A, j\in B} f(r_{ij}) = \sum_{i\in A, j\in B} \frac{1-(\frac{r_{ij}}{r_0})^m}{1-(\frac{r_{ij}}{r_0})^n} \tag{1}$$

where i and j are indices of atoms of group A and B to be considered, r_{ij} is the distance between atoms i and j. The groups of A and B are donor group NH and the acceptor group CO for BC and are identical group of the heavy atoms of side-chain groups. The reference distance, r_0, are chosen as 2.5Å and 4.5Å for BC and SC respectively on the basis of the geometrical criteria of the hydrogen bonds and van-der-Waals interaction. The exponent values, m and n, are chosen as 30 and 60 for a sharp switching function.

SI-2 Free Energy Surface for MET system

SI-3 Absorption Spectra of all structures of NC and MET systems

The calculated absorption spectra of all configurations on the different states are shown in Figures S2 to S4. The left and right panels of Figure S2 show the absorption spectra of six configurations on different states including transition states X and Y in the gas phase

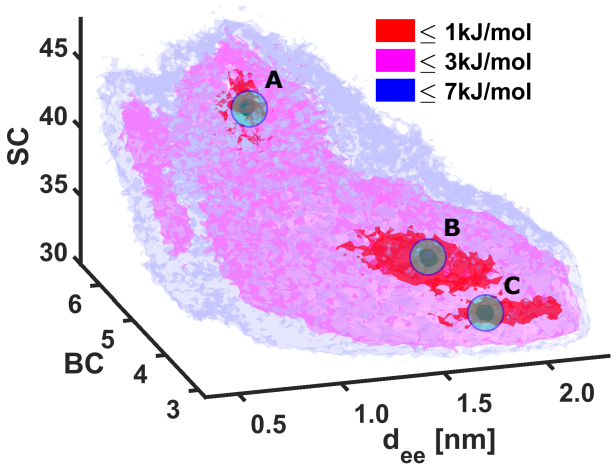

Figure S1: 3D free energy surface along d_{ee}, BC and SC for MET system.

and the protein environment, respectively. While Figure S3 shows the absorption spectra of the configurations on the states of MET system in the gas phase, Figure S4 shows the spectra for the same MET configurations in the protein environment.

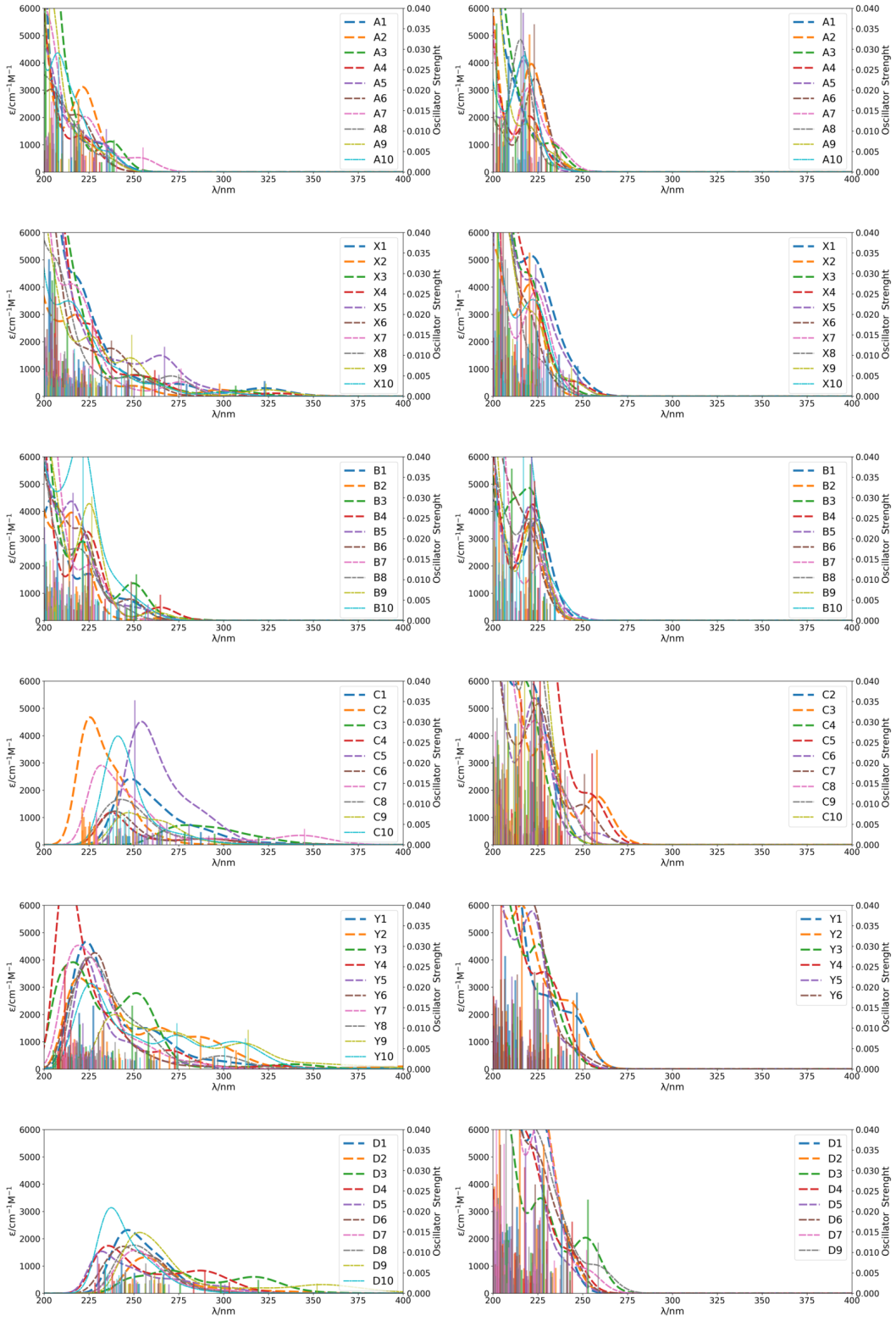

Figure S2: Calculated absorption spectra of NC configurations in the gas phase (left) and the protein environment (right)

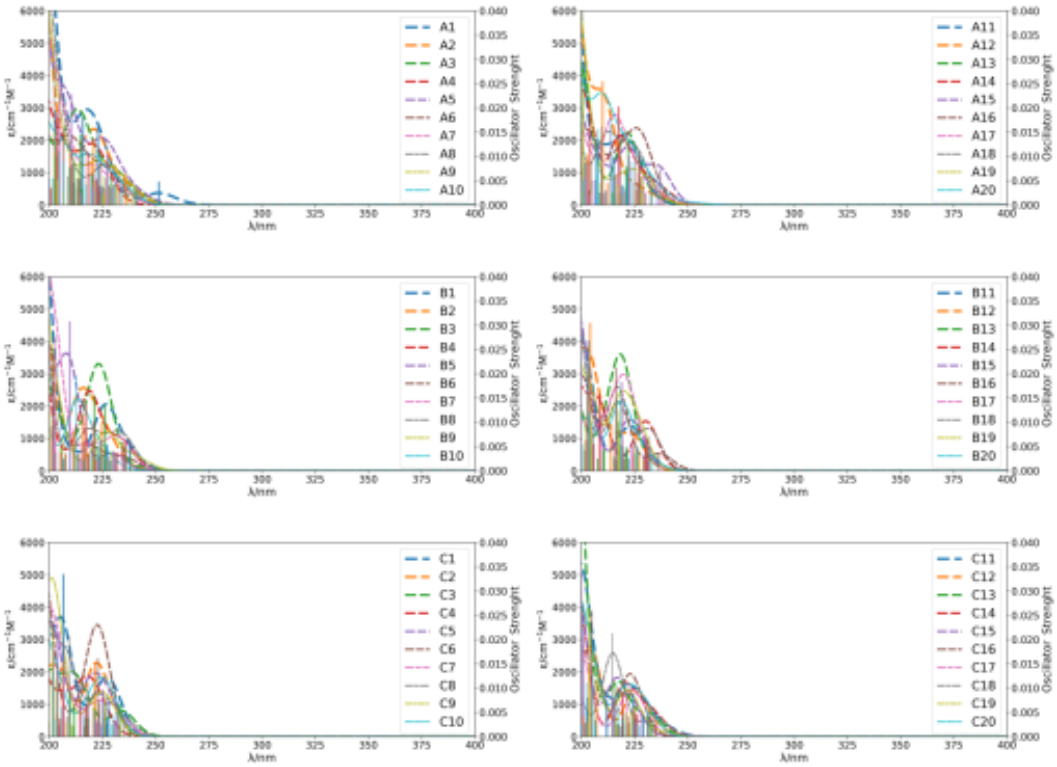

Figure S3: Calculated absorption spectra of MET configurations in the gas phase

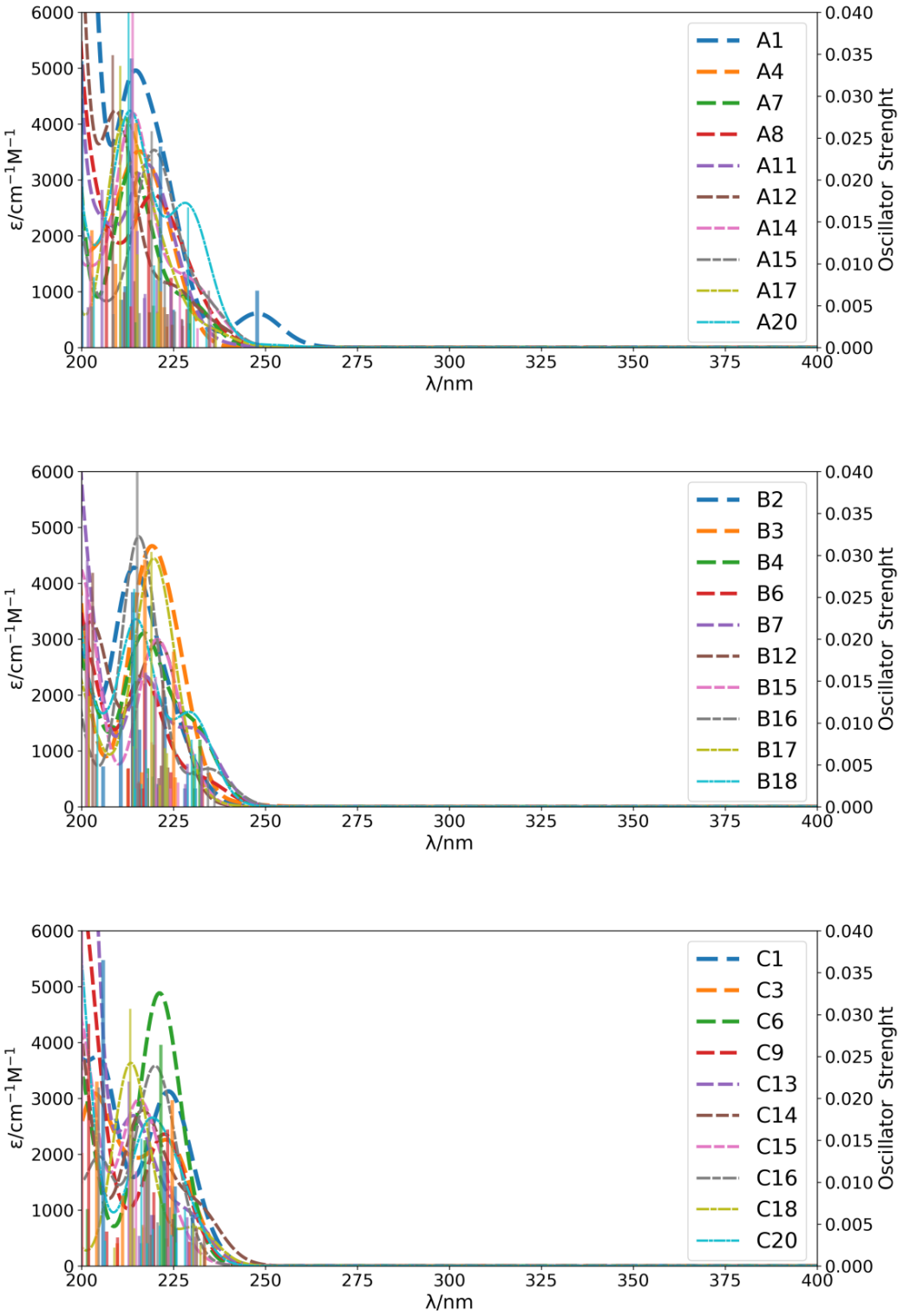

Figure S4: Calculated absorption spectra of MET configurations in the protein environment

SI-4 Molecular Orbitals

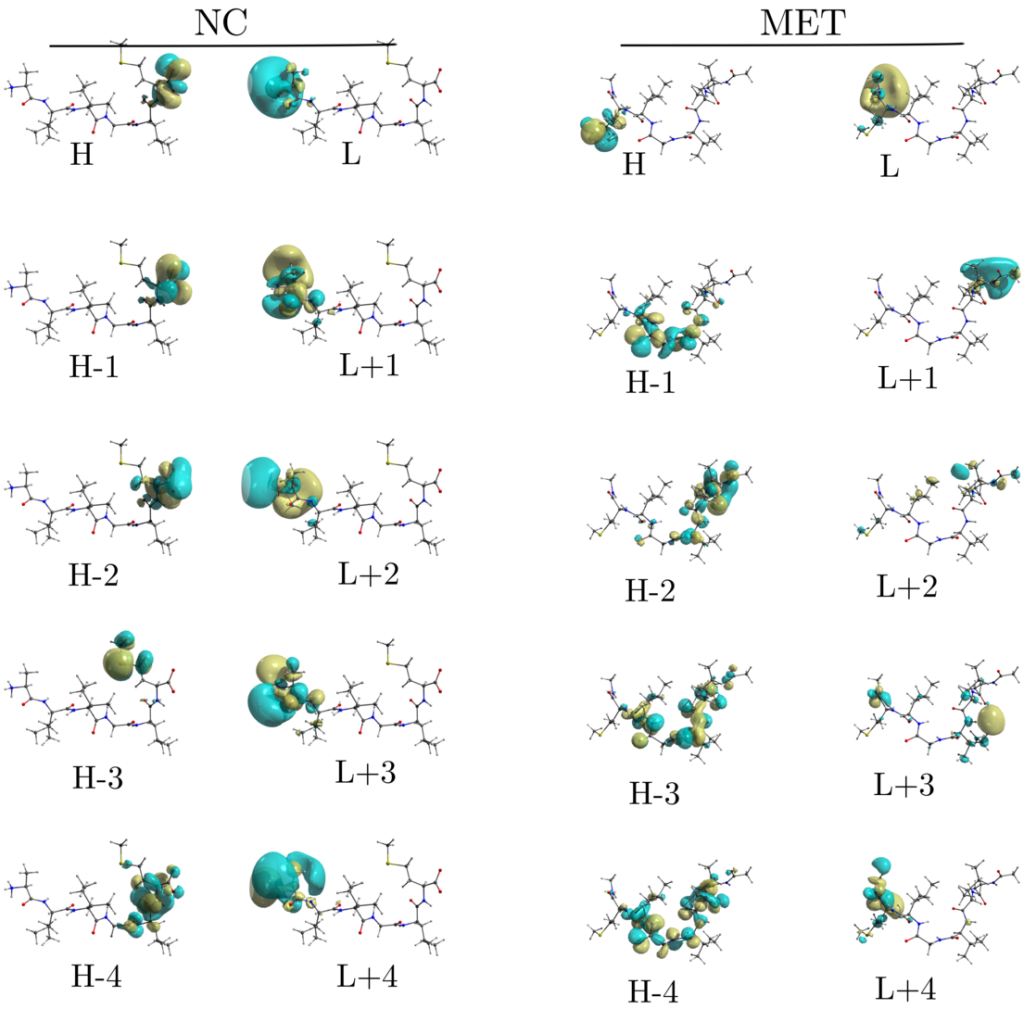

Figure S5: Kohn-Sham frontier molecular orbitals for a selected zwitterionic (left) and methyl-capped (right) structures.

SI-5 Characterization of the Low energy excitation on the configurations of state C of NC system

Table S1: The characterization of the excited states for the configurations on state C of NC system in the protein environment. The first column is the configuration name, second column is for the type of the environment, third column is the index of the excited states, fourth 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

Conf.	Env.	state	$\lvert\Delta\mu\rvert$	$\Delta\mathbf{r}$ (Å)
C2	PCM	S1	0.77	0.79
C2	PCM	S1	1.05	0.99
C4	PCM	S8	0.57	0.71
C5	PCM	S1	0.36	0.44
C6	PCM	S1	2.8	2.32
C7	PCM	S1	3.06	2.49
C8	PCM	S1	0.84	0.77
C9	PCM	S2	0.75	0.71
C10	PCM	S4	0.8	0.81

Table S2: The characterization of the excited states for the configurations of A3, A10, C5 and C7. The first column is the configuration name, second column is for the type of the environment, third column is the index of the excited states, the forth and fifth are the corresponding wavelength and 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.

Env.	Conf.	λ(nm)	state	Perc.(%)	f_I	$\|\Delta\mu\|$ (eÅ)	$\Delta\mathbf{r}$ (Å)
GAS	A3	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
	A10	237.5	S2	68.2	0.006	2.98	2.14
			S1	26.1	0.002	1.19	1.28
	C3	324.6	S47	67.0	0.002	7.36	4.88
			S43	16.3	0.001	9.34	5.82
	C5	289.7	S63	20.7	0.005	6.89	4.71
			S53	15.8	0.002	7.29	5.28
			S56	12.6	0.001	7.79	5.34
			S54	11.6	0.001	8.64	6.31
			S52	9.4	0.002	5.55	4.06
	C7	345.4	S14	90.3	0.004	5.69	3.62
PCM	A3	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
	A10	237.1	S1	95.7	0.003	1.03	1.17
	C3	257.8	S2	99.9	0.026	1.05	0.99
	C5	256.7	S1	98.3	0.022	0.36	0.44
	C7	251.0	S1	94.2	0.017	3.06	2.49

Table S3: ACE system.

Env.	Conf.	λ(nm)	state	Perc.(%)	f_I	$\|\Delta\mu\|$ (eÅ)	$\Delta\mathbf{r}$ (Å)
GAS	C3	243.1	S4	51.2	0.002	1.62	1.65
			S3	44.7	0.001	3.36	2.79
	C5	273.1	S1	97.7	0.002	3.35	2.69
PCM	C3	222.2	S2	66.5	0.034	0.91	1.08
			S4	10.0	0.005	0.21	0.28
			S3	9.6	0.005	0.60	0.77
			S5	6.9	0.003	0.30	0.39
	C5	225.0	S5	37.3	0.031	1.10	1.11
			S4	20.5	0.010	0.41	0.30
			S3	17.9	0.008	0.30	0.40
			S6	13.3	0.012	0.30	0.40

SI-6 Correlation between Structural Parameters and E_1 (λ)

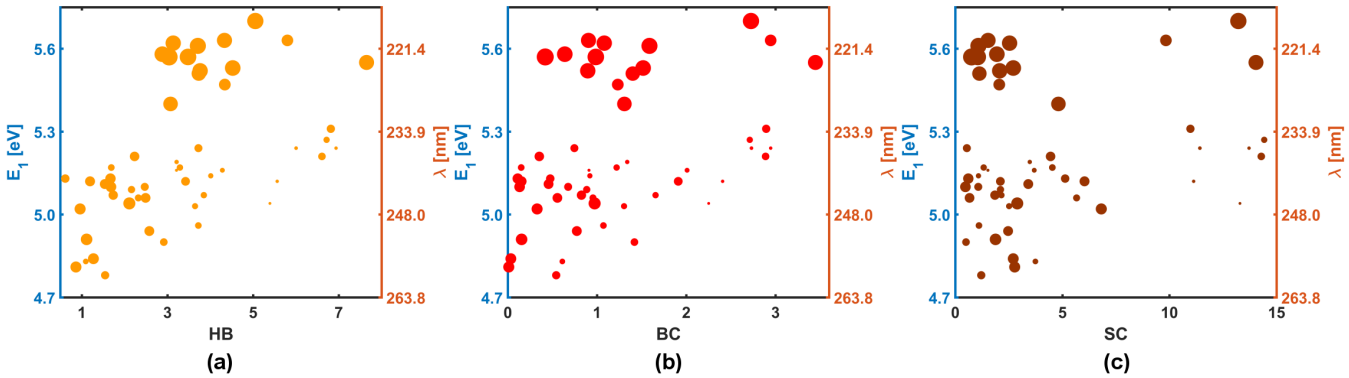

Figure S6: Scatter plot between (a) HB and E_1 (λ), (b) BC and E_1 (λ) and (C) SC and E_1 (λ) for NC system in the protein environment.

Table S4: Pearson's and Spearman's correlation coefficients between HB and E_1, BC and E_1, and SC and E_1 for NC system in gas phase.

	Pearson	Spearman
r(HB, E_1)	0.7366	0.6852
r(BC, E_1)	0.6290	0.4945
r(SC, E_1)	0.5388	0.2551

To illustrate the extent of the reduction of the correlation made by the solvent effects, we used Pearson's and Spearman's correlation coefficients. Table S5 shows the values of Pearson's and Spearman's correlation coefficients between the structural parameters of HB, BC and SC and E_1.

The weak correlation of the ground stats and excited sates with the conformational fluctuations for MET configurations might provide the clue to the relationship between the absorption spectra and the structural parameters. The values of E_{exc} and λ in the absorption spectra of the forty configurations of MET system in Figure S3 have been calculated to investigate the correlation of these quantities with the structural parameters. In Figure S7,

Table S5: Pearson's and Spearman's correlation coefficients between HB and E_1, BC and E_1, and SC and E_1 for NC system in protein environment.

	Pearson	Spearman
r(HB, E_1)	0.4976	0.5882
r(BC, E_1)	0.3849	0.4301
r(SC, E_1)	0.1697	0.1132

we showed 2D scatter plots of E_{exc} and λ versus BC and SC, where E_{exc} is plotted on the left y-axis and λ is plotted on the right y-axis. The scatter plot for BC in the left panel

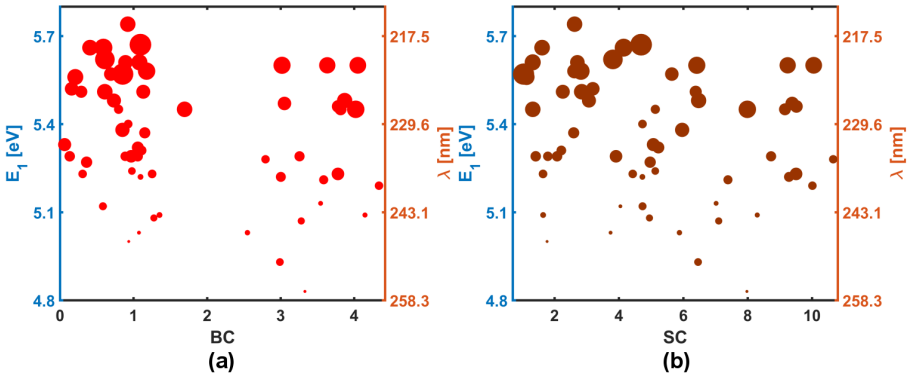

Figure S7: Scatter plot between (a) Bc and E_{exc} (λ) and (b) SC and E_{exc} (λ) for MET system in gas phase

of Figure S7 shows basically no correlation between the E_{exc} and BC. Besides four points corresponding to the relatively large E_{exc}, most of configurations have the similar E_{exc} of around 5.5 (eV)with the corresponding λ of about 225.5 (nm). Similarly, the right panel of Figure S7 for the SC scatter plot is likely to display no correlation, but the four points with large E_{exc} corresponds to the large values of SC, which would provide a bit larger correlation than BC. This correlation can also be quantified by Pearson's and Spearman's correlation coefficients represented in Table S6. Comparing to the correlation coefficients of NC configurations, both cases of BC and SC have the relatively small values of about 0.1 and 0.2, respectively, and the correlation coefficients between SC and E_{exc} are larger than ones between BC and E_{exc}, as expected.

Table S6: Pearson's and Spearman's correlation coefficients between pairs of BC and E_{exc}, and SC and E_{exc} for MET system in gas phase.

	Pearson's	Spearman's
r(BC, E_{exc})	-0.2344	-0.2687
r(SC, E_{exc})	-0.1987	-0.2314

The electronic structure of MET configurations are not affected a lot by the introduction of the protein environment, but it's still worthwhile probing the influence of the protein environment on the correlation between the $E_{exc}(\lambda)$ and structural parameters because the protein environment should be considered due to its importance as a model of the real fibrils' environment around the monomer. The absorption spectra of MET system in the protein environment in the right panel of Figure S3 show the increased oscillator strength with respect to the gas phase in the left panel of Figure S3, but there is no big change in the positions of the first peaks. The calculated E_{exc} and λ are plotted with the BC and SC in 2D scatter plots depicted in Figure S8. The number of points with the large values of E_{exc} apparently decreased from four to two on both left and right panels of Figure S8 and less counted number of the points in the scatter plots implies the overlap of these points. Table S7 shows the Pearson's and Spearman's correlation coefficient of the scatter plots

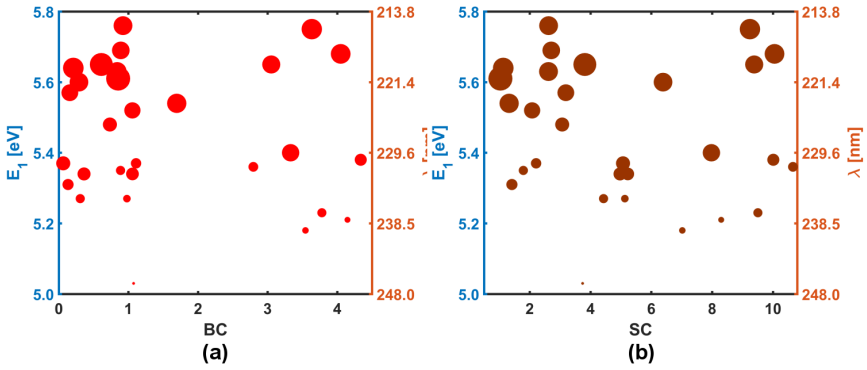

Figure S8: Scatter plot between (a) BC and E_{exc} (λ) and (b) SC and E_{exc} (λ) for MET system in the protein environment.

in Figure S8. The absolute values of the correlation coefficients for BC and SC are around

zero which shows no correlation.

Table S7: Pearson's and Spearman's correlation coefficients between pairs of BC and E_{exc}, and SC and E_{exc} for MET system in protein environment.

	Pearson's	Spearman's
r(BC, E_{exc})	-0.1116	-0.1062
r(SC, E_{exc})	-0.1383	-0.1594

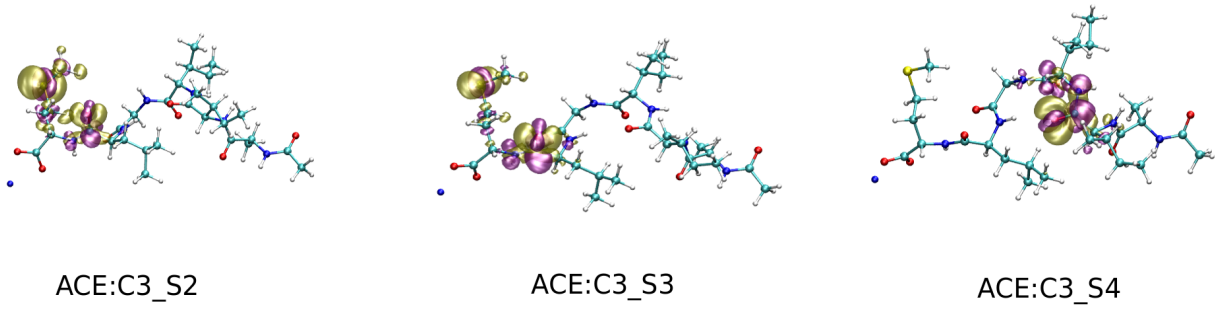

Figure S9: so-surface plot of charge difference between ground and excited states for acetylated systems. The violet and yellow colors represent charge increment and depletion. Isovalues are set to ±0.002 a.u.

SI-7 Experimental optical spectra

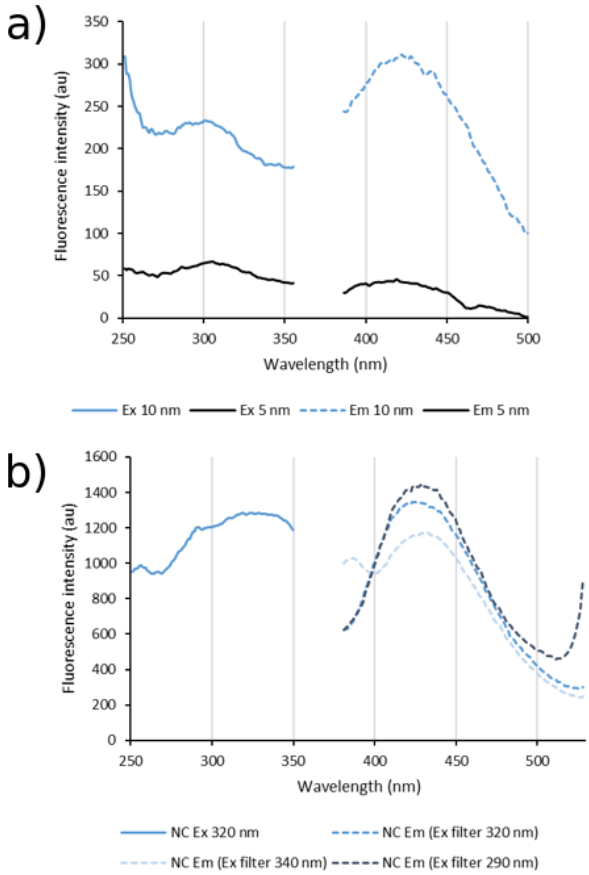

Figure S10: (a) Comparison of 5 nm excitation slit to 10 nm excitation slit for ACE aggregated in phosphate buffered saline. For excitation spectrum the emission was collected at 420 nm where the excitation slit was 10 nm (blue) or 5 nm (black) and the emission slit was 20 nm. For the emission spectra the excitation was set at 305 nm where the excitation slit was 10 nm (blue) or 5 nm (black) and the emission slit was 20 nm. Reducing the excitation slit lead to great reduction in signal. (b) Emission spectra of NC collected with different excitation filters. For excitation spectrum the emission was collected at 420 nm where the excitation slit was 10 nm and the emission slit was 20 nm. For the emission spectrum the excitation was set at 340 nm (light blue), 320 nm (mid blue) and 290 nm (dark blue) where the excitation slit was 10 nm and the emission slit was 20 nm. Reducing the excitation slit lead to great reduction in signal.

References

(S1) Jong, K.; Grisanti, L.; Hassanali, A. Hydrogen Bond Networks and Hydrophobic Effects in the Amyloid Aβ_{30-35} Chain in Water: A Molecular Dynamics Study. *Journal of Chemical Information and Modeling* **2017**, *57*, 1548–1562, PMID: 28603985.