Electronic Supplementary Information (ESI)

Monitoring Early-Stage β-Amyloid Dimer Aggregation by Histidine Site-Specific Two-Dimensional Infrared Spectroscopy in a Simulation Study

Sompriya Chatterjee[¶,†], *Yeonsig Nam*[¶,§,†], *Abbas Salimi*[¶,†], *and Jin Yong Lee*[¶,*]

[¶] *Department of Chemistry, Sungkyunkwan University, Suwon 440-746, Korea*

[§] *Department of Chemistry, University of California, Irvine, California 92697-2025, United States*

[†]*These authors contributed equally to this work*

**Corresponding author: jinylee@skku.edu (J. Y. Lee);*

Contents

1 Binding free energy calculations

The binding free energy between the monomers that form dimers was determined by the Molecular Mechanics-Poisson Boltzmann surface area (MM-PBSA) approach using the following equation:

$$\Delta E_{binding} = G_{dimeric\ complex} - (G_{monomer\ 1} + G_{monomer\ 2}) \quad \text{(S1)}$$

Moreover, for each individual part, the free energy can be given by:

$$G_i = < E_{MM} > - TS + < G_{solv} > \quad \text{(S2)}$$

where G_i indicates the total free energy. [1] $< E_{MM} >$ is the average molecular mechanics potential energy, containing energy of both bonded and non-bonded terms. The bonded interactions consist of the angle, bond, and dihedral association while the van der Waals and electrostatic connections produce the non-bonded part. In addition, $< G_{solv} >$ implies the sum of polar solvation energy and the non-polar solvation component, and TS denotes conformational entropy.[1] The non-polar solvation part is estimated by solvent accessible surface area (SASA). As the binding energy calculated here is the relative binding free energy, the entropic contribution of Aβs was ignored, which is in agreement with a number of earlier theoretical analyses.[2,3]

Comparison of the binding free energy of the protonated (πππ:πππ) system with tautomeric dimers (δδδ:δδδ and εεε:εεε) during aggregation has not yet been determined. The van der Waals and non-polar terms (SASA) were negative in all dimers and favored complex formation (Table S1). On the other side, $\Delta E_{binding}$ of the δδδ:δδδ, and εεε:εεε systems included unfavorable contributions of electrostatic energy. However, within πππ:πππ, the electrostatic contribution was one of the main contributing parameter to stabilize the dimeric complex. In contrast, the positive value of ΔG_{polar}

in πππ:πππ indicated its destabilizing effect. In the other two dimers, δδδ:δδδ and εεε:εεε, we observed unfavorable and favorable contributions of polar solvation energy.

2 Frequency calculations

Before calculating 2D IR spectra, we determined vibrational frequency distributions within the dimer (Figure S3). Numerous parametrization schemes have been utilized for the vibrational mode in biomolecules. To obtain the amide-I frequency, the CHO4 parameterization, which is based on an extension of the vibration frequency as a linear combination of the electrostatic potential calculated at the four coordinates of C, O, N, and H of each amide bond, was used.[4,5] A previous work by Falvo and colleagues exhibited excellent correlation between theoretically and experimentally determined residual frequency fluctuations in Aβ fibril by the CHO4 parameterization.[6] The vibration frequency of the n^{th} protein bond positioned between the *r* and *r* + 1 residues was generated as:

$$\hbar\omega_n = \hbar\omega_0 + \sum_{s=C,O,N,H} l_s \emptyset_{n,s}(t). \quad \text{(S3)}$$

The sum is carried out over the C, N, H, and O atoms of the protein bond *n*. Similar to our monomeric work,[7] we fixed the central band ($\hbar\omega_0$) to 1600 cm^{-1}. The l_s values are given by l_O = 0.00160*e*, l_C = −0.00554*e*, l_N = 0.00479*e*, and l_H = −0.00086*e*, where *e* is the electronic charge. The electrostatic potential $\phi_{n,s}$ was computed at the $r_{n,s}$(t) coordinate of the atom *s* of the n^{th} protein bond as the sum of the backbones and side chains contribution:

$$\emptyset_{n,s} = \emptyset_{n,s}^{backbones}(t) + \emptyset_{n,s}^{side\ chains}(t), \quad \text{(S4)}$$

$$\emptyset_{n,s}^{P}(t) = \frac{1}{4\pi\varepsilon_0} \sum_{i\,\epsilon\,P} \frac{q_i}{|r_i(t) - r_{n,s}(t)|}, \quad \text{(P = backbones or side chains)} \quad \text{(S5)}$$

where $\phi_{n,s}^{backbones}$(t) and $\phi_{n,s}^{side\ chains}$(t) correspond to the electrostatic potential created by the Aβ dimer backbone and side chain atoms. Furthermore, the electrostatic potential contribution is obtained by the sum of point charges, q_i, over distance $\left|r_i(t) - r_{n,s}(t)\right|$ and ε_0 for all atoms (i) in the backbones and side chains. Similar to our previous research on the Aβ40 monomer,[7] the q_i of dimeric side chain atoms were calculated using the ff99SB force field. An earlier evidence suggested that the ff99SB parameter set supplied a reasonable parameter set for biomolecular simulations.[8] For backbone (C, O, N, H, C_α, and H_α) contributions, we followed the study of Ham *et al*.[4,5] who considered the values for q_{C_α}, q_{H_α}, q_C, q_O, q_N, and q_H as 0, 0, 0.419, −0.871, 0.793, and −0.341, respectively. Ham *et al*. compared ab initio and map frequency shift of peptides and found a good agreement between them using these values.

3 Supplementary figures

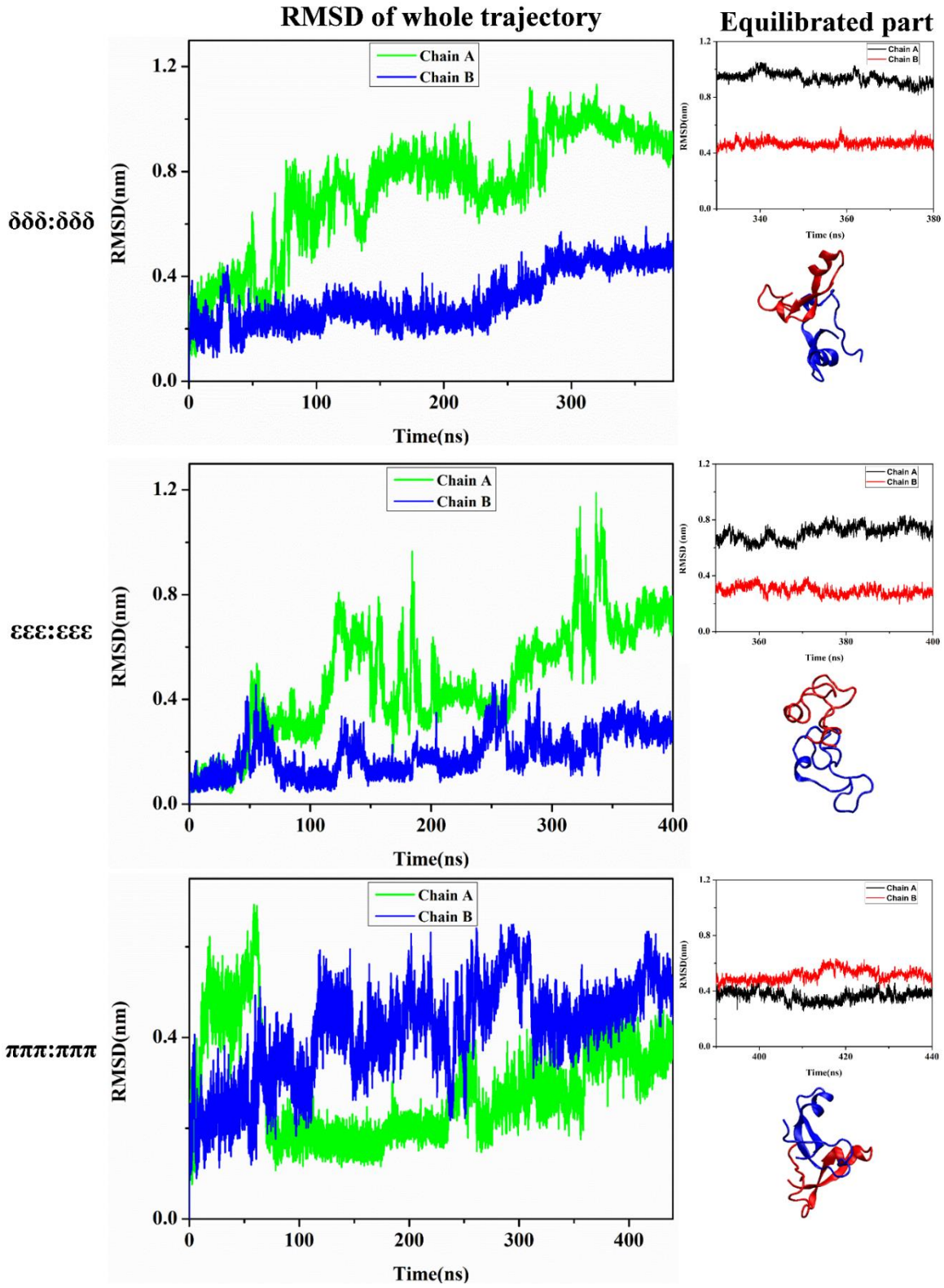

Figure S1. Root mean square deviation (RMSD) of residual backbone of δδδ:δδδ, εεε:εεε, and πππ:πππ dimers in total and converged part of each trajectory with initial structures.

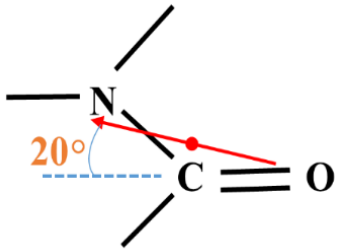

Figure S2. Schematic drawing of the location (red circle) and direction (red arrow) of the transition dipole used in this study.

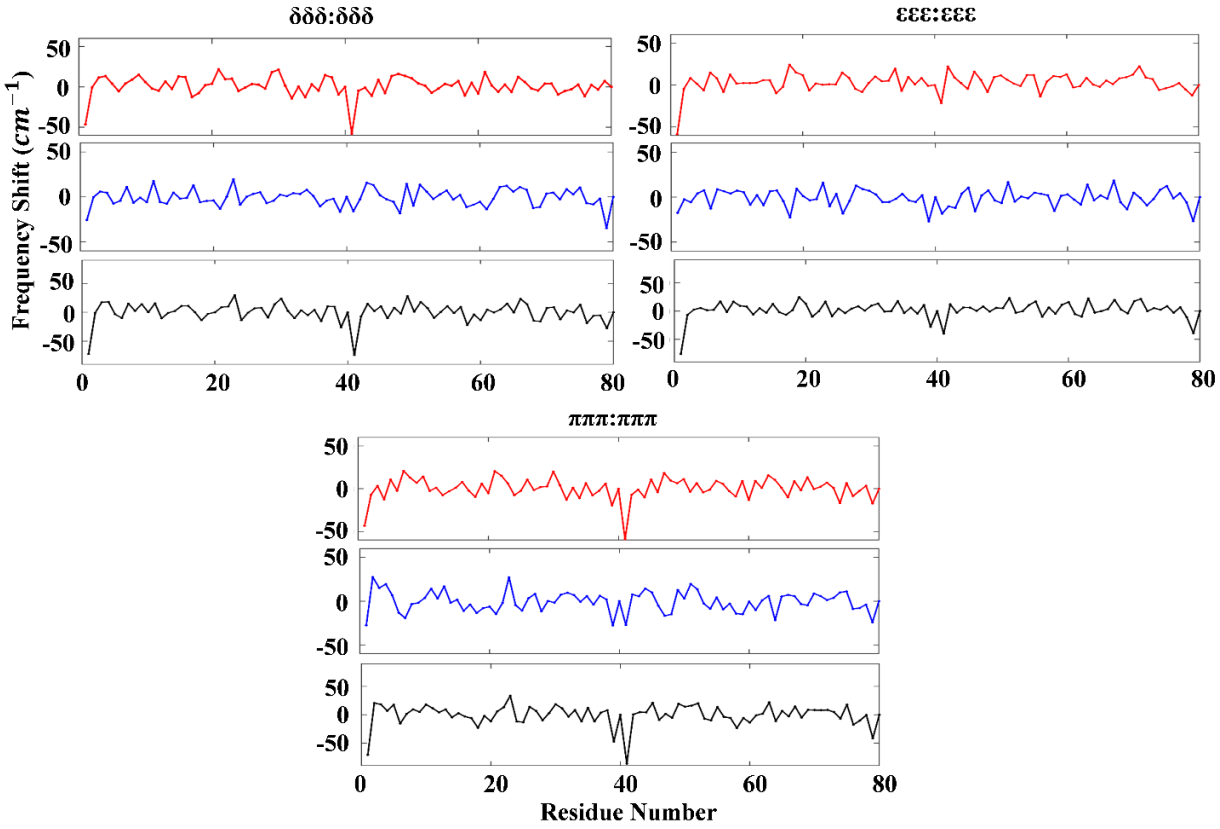

Figure S3. The average frequency shift, $\langle \hbar\omega_0 - \hbar\omega_n \rangle$, of amide-I vibration induced by the electrostatic potential of the backbone, and side chains. The black, red, and blue lines refer to the

average frequency shifts of dimers due to the total (backbone + side chain) protein, backbone, and side chains, respectively.

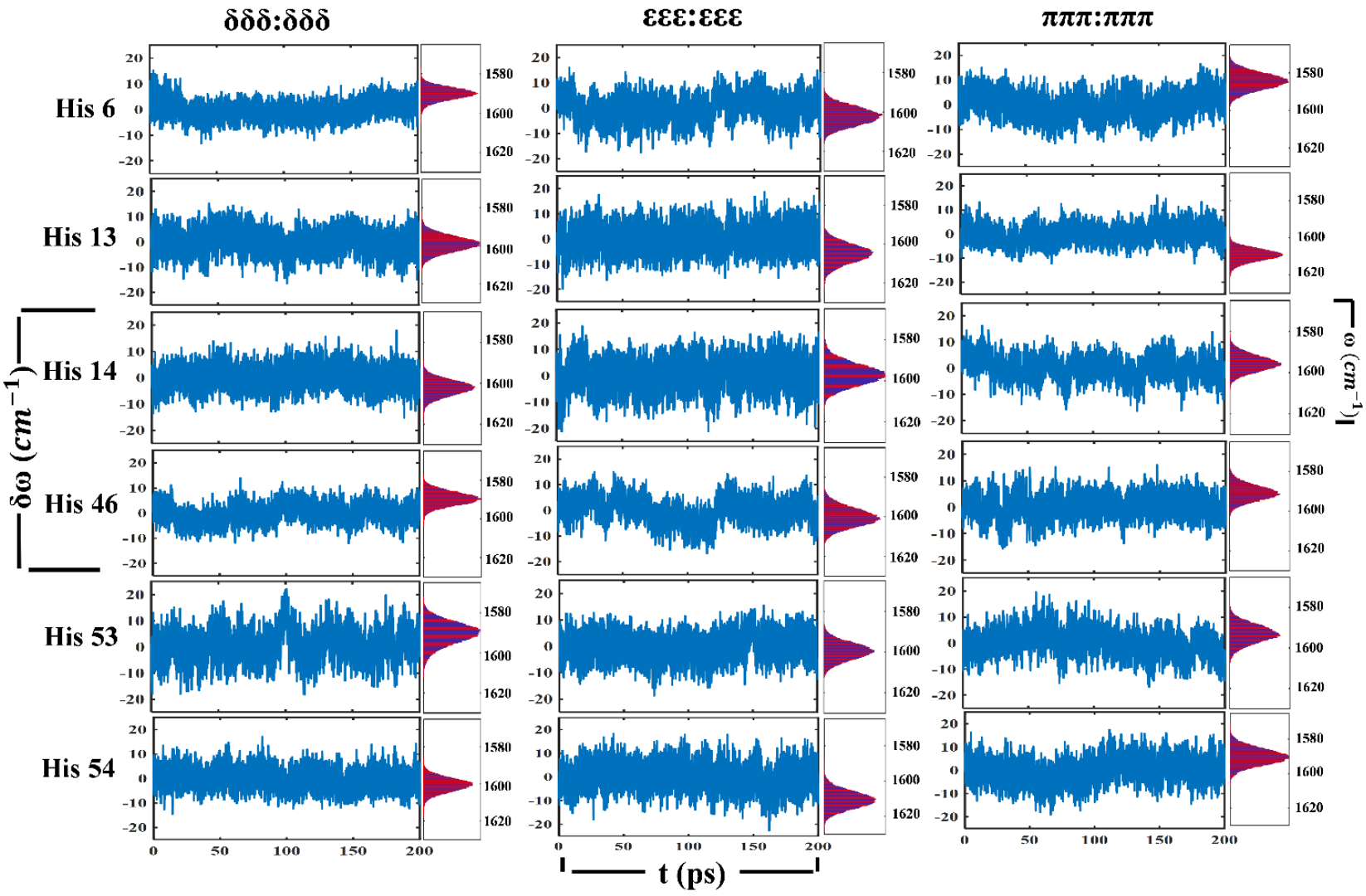

Figure S4. Frequency deviations and distributions of histidine in Aβ40 dimers. Frequency and time units are given in cm^{-1} and ps, respectively.

4 Supplementary tables

Table S1. Average Van der Waals, electrostatic, solvent accessible surface area, polar solvation, and binding free energy for each dimer. (Parentheses indicate the standard error) (Energy units are in kJ/mol).

Dimers	ΔE_{vdw}	ΔE_{elec}	ΔG_{polar}	ΔG_{sasa}	$\Delta E_{binding}$
δδδ : δδδ	-152.456 (2.114)	15.322 (1.884)	21.240 (3.909)	-17.516 (0.248)	-133.373 (1.217)
εεε : εεε	-105.762 (1.588)	111.535 (0.635)	-152.205 (1.276)	-12.072 (0.186)	-158.622 (1.443)
πππ : πππ	-15.415 (0.707)	-6.083 (0.319)	22.587 (1.043)	-1.657 (0.078)	-0.540 (0.649)

Table S2. Parameters obtained from fitting the FFCF. Here, τ_1, τ_2, τ_3, and γ are given in ps^{-1}.

δδδ: δδδ	a_1	γ	τ_1	a_2	τ_2	a_3	τ_3
His 6	0.2836	31.3921	19.0152	0.3415	9.5244	0.3114	0.0824
His 13	0.4246	13.8686	34.1696	0.3307	16.6173	0.2487	0.1255
His 14	0.2957	27.0168	18.2420	0.5256	19.1260	0.1252	0.0878
His 46	0.3982	33.0579	18.5208	0.4002	9.4549	0.3764	0.0865
His 53	0.3360	25.4656	18.2159	0.5618	10.5719	0.3365	0.2275
His 54	0.5826	16.3368	8.4836	0.2832	22.4910	0.1497	0.2503
εεε:εεε							
His 6	0.4676	25.7328	27.4524	0.3829	6.7148	0.3679	0.1853
His 13	0.4925	12.5483	4.6523	0.2350	18.4312	0.2434	2.1248
His 14	0.4190	44.8515	26.6529	0.3773	10.1619	0.2018	0.2830
His 46	0.2611	17.6551	28.0602	0.3199	12.1556	0.5309	0.0980
His 53	0.5874	31.4824	16.5080	0.3321	11.0780	0.2064	0.2075
His 54	0.4155	21.6772	17.6728	0.3862	14.0373	0.2100	0.2752
πππ:πππ							
His 6	0.2794	16.3661	18.8987	0.3794	14.8039	0.3328	0.1000
His 13	0.4358	25.5970	24.1166	0.3913	11.1287	0.2574	0.2135
His 14	0.2634	75.5993	26.2464	0.4019	5.5404	0.3784	0.1875
His 46	0.4988	25.3877	22.6755	0.3986	12.9756	0.3081	0.3814
His 53	0.4624	39.4040	25.8612	0.4196	5.8628	0.3361	0.1010
His 54	0.4653	35.3141	18.9580	0.3407	11.1486	0.2892	0.1593

Table S3. Comparison of the average distances (unit: Å) between histidine and other residues in monomer[7] and dimer. The lengths were determined by calculating the distance between the dipole center every time step and averaged over 200 ps.

	His 6	**His 46**	**His 13**	**His 53**	**His 14**	**His 54**
εεε[7]	12.3		15.6		14.8	
εεε:εεε	18.2	13.6	20.4	14.8	18.5	13.6
δδδ[7]	11.8		13.8		16.3	
δδδ:δδδ	14.4	13.5	16.6	18.2	16.1	19.9
πππ[7]	11.8		15.3		18.3	
πππ:πππ	12.8	11.9	16.2	19.4	16.2	17.1

5 References

1 R. Kumari, R. Kumar and A. Lynn, *Journal of Chemical Information and Modeling*, 2014, **54**, 1951–1962.

2 Y. Sun, Z. Qian and G. Wei, *Physical Chemistry Chemical Physics*, 2016, **18**, 12582–12591.

3 Y. Sun, W. Xi and G. Wei, *The Journal of Physical Chemistry B*, 2015, **119**, 2786–2794.

4 J.-H. Choi, S. Ham and M. Cho, *The Journal of Physical Chemistry B*, 2003, **107**, 9132–9138.

5 S. Ham, S. Cha, J.-H. Choi and M. Cho, *The Journal of Chemical Physics*, 2003, **119**, 1451–1461.

6 C. Falvo, W. Zhuang, Y. S. Kim, P. H. Axelsen, R. M. Hochstrasser and S. Mukamel, *The Journal of Physical Chemistry B*, 2012, **116**, 3322–3330.

7 Y. Nam, M. Kalathingal, S. Saito and J. Y. Lee, *Biophysical Journal*, 2020, **119**, 831–842.

8 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Structure, Function, and Bioinformatics*, 2006, **65**, 712–725.