

# Tautomeric Effect of Histidine on $\beta$ -Sheet Formation of Amyloid Beta 1–40: 2D-IR Simulations

Yeonsig Nam,<sup>1,2</sup> Mahroof Kalathingal,<sup>3</sup> Shinji Saito,<sup>2,3,\*</sup> and Jin Yong Lee<sup>1,\*</sup>

<sup>1</sup>Department of Chemistry, Sungkyunkwan University, Suwon, Korea; <sup>2</sup>Institute for Molecular Science, Myodaiji, Okazaki, Japan; and <sup>3</sup>The Graduate University for Advanced Studies, Myodaiji, Okazaki, Japan

**ABSTRACT** Histidine state (protonated or  $\delta$  or  $\epsilon$  tautomer) has been considered the origin of abnormal misfolding and aggregation of  $\beta$ -amyloid ( $A\beta$ ). Our previous studies reported that the  $\delta\delta\delta$  isomer of  $A\beta$  (1–40) has a greater propensity for  $\beta$ -sheet conformation compared to other isomers. However, direct proof of the tautomeric effect has not been reported. In this context, we calculated histidine site-specific two-dimensional infrared spectroscopy of the  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  (all protonated histidine) systems within the framework of classical molecular dynamics simulations aiming at connecting our previous results with the current experimental observations. Our results showed that  $\beta$ -sheet formation is favored for the  $\delta\delta\delta$  and  $\pi\pi\pi$  tautomers compared with the  $\epsilon\epsilon\epsilon$  tautomer, consistent with our previous studies. This result was further supported by contact map analyses and the strength of dipole coupling between the amide-I bonds of each residue. The two-dimensional infrared diagonal trace for each tautomer included three distinctive spectrally resolvable peaks near 1680, 1686, and 1693  $\text{cm}^{-1}$ , as was also observed for histidine dipeptides. However, the peak positions at His6, His13, and His14 did not show a consensus trend with the histidine or protonation state but were instead affected by the presence of surrounding hydrogen bonds. Our study provides a deeper insight into the influence of tautomerism and protonation of histidine residues in  $A\beta$  (1–40) on amyloid misfolding and provides a connection between our previous simulations and experimental observations.

**SIGNIFICANCE** The amyloid hypothesis of Alzheimer's disease is well recognized. However, there is no answer why amyloids aggregate. Recently, histidine tautomerism hypothesis has been proposed for the aggregation of amyloids by our group. Here, we tested 2DIR spectra to distinguish the histidine tautomers in amyloid beta peptide. The 2DIR diagonal peak of each tautomer shows three distinctive peaks; however, the peak positions at His6, His13, and His14 do not show a consensus trend with the histidine/protonation state, but rather, are affected by the presence of surrounding hydrogen bonding and donor-acceptor pairs. Our study provides a deeper insight into the effect of tautomerism/protonation of histidine residues in  $A\beta$  (1–40) on amyloid misfolding and provides a bridge between our previous simulations and experimental observations.

## INTRODUCTION

Alzheimer's disease (AD) is one of the most common forms of dementia, affecting more than 47 million people worldwide, and results in problems with language, disorientation, and mood swings. AD is generally thought to result from entanglement of  $\tau$  proteins and/or aggregation and accumulation of senile  $\beta$ -amyloid peptide plaques ( $A\beta$ ) (1,2). The  $A\beta$  hypothesis for the origin of AD is widely accepted, and  $A\beta$  is assumed to play an important role in AD. Even small perturbations of the physiological environment can in-

fluence the chemical and dynamic properties of peptides and destroy the balance of production and clearance of  $A\beta$ . Many studies have focused on the influence of external factors (3–6) such as pH changes, amyloid precursor protein mutation, and the presence of metal ions on the aggregation of  $A\beta$ . In particular, slightly acidic conditions can promote formation of  $\beta$ -sheets and aggregation of  $A\beta$ , which may be one reason why the brains of AD patients are relatively acidic compared with the brains of normal people (7).

More recently, a tautomer hypothesis for the origin of AD has been suggested (8,9): the tautomeric state of the histidine side chain (Fig. 1) plays a crucial role in  $A\beta$  misfolding, which implies that  $A\beta$  aggregation may arise intrinsically rather than as a result of external factors. Histidine residues are critical to the stability of the fibril structure

Submitted February 3, 2020, and accepted for publication July 7, 2020.

\*Correspondence: [shinji@ims.ac.jp](mailto:shinji@ims.ac.jp) or [jinyilee@skku.edu](mailto:jinyilee@skku.edu)

Editor: Elsa Yan.

<https://doi.org/10.1016/j.bpj.2020.07.009>

© 2020 Biophysical Society.



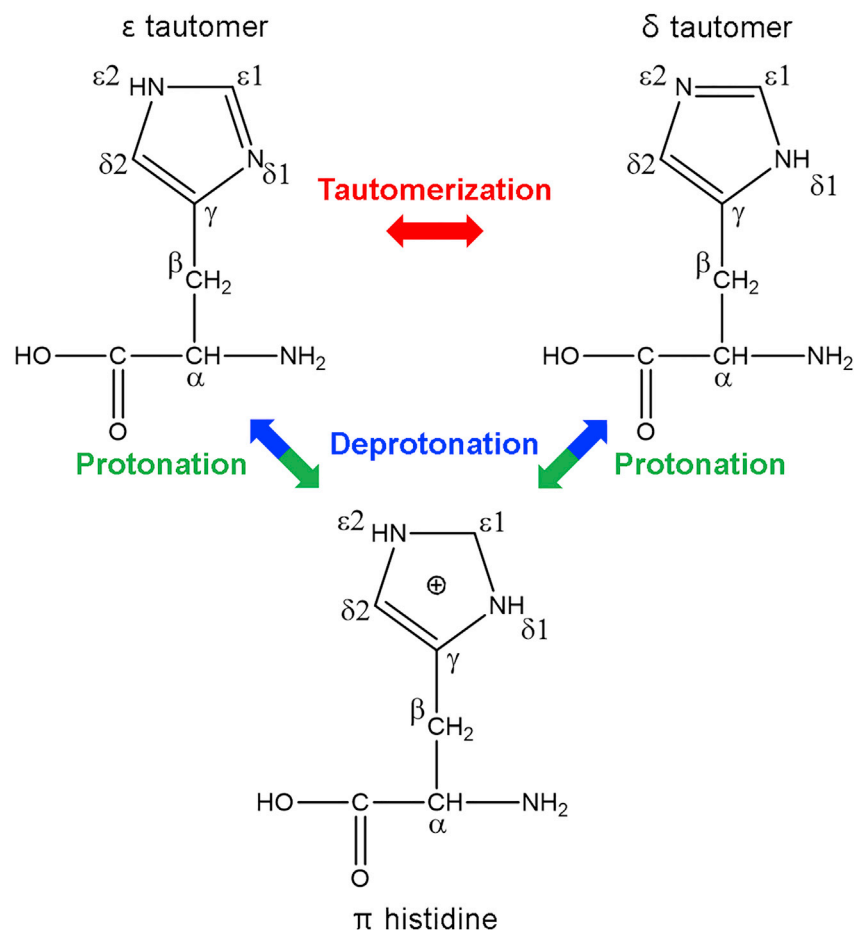


FIGURE 1 Two tautomeric ( $\delta$  and  $\epsilon$  histidine) and protonated ( $\pi$  histidine) forms of histidine. To see this figure in color, go online.

(10). As an example, the presence of three protonated histidine residues reduces the electrostatic repulsion between the two terminal regions of the A $\beta$  protein, leading to greater probability of A $\beta$  (1–42) misfolding and aggregation (11).

Shi et al. (9) have used molecular dynamics (MD) simulations to demonstrate the tautomer effects of the A $\beta$  (1–40) peptide, which, when containing the His6( $\delta$ )-His13( $\delta$ )-His14( $\delta$ ) isomer (the so-called  $\delta\delta\delta$  isomer), leads to the formation of greater amounts of antiparallel  $\beta$ -sheets between the central hydrophobic center (CHC) and C-terminus compared with the other seven isomers ( $\delta\delta\epsilon$ ,  $\delta\epsilon\delta$ ,  $\delta\epsilon\epsilon$ ,  $\epsilon\epsilon\epsilon$ ,  $\epsilon\epsilon\delta$ ,  $\epsilon\delta\epsilon$ ,  $\epsilon\delta\delta$ ). Our previous simulation studies regarding homodimers (12) and pentamers (13) confirmed such trends. However, the effect of tautomeric state of histidine within A $\beta$  requires clarification. Some studies have focused on distinguishing the tautomeric forms of the histidine imidazole ring. Ghosh et al. (14) applied linear infrared (IR) spectroscopy to a histidine dipeptide system and reported that the  $\delta$  form had a higher peak position compared to the  $\epsilon$  form. In addition, Cheng et al. (15) and Vila et al. (16) postulated that the fraction of tautomeric forms within the imidazole ring in histidine dipeptides and tripeptides can be distin-

guished by applying NMR spectroscopy. However, the usefulness of NMR spectroscopy may be limited when tautomeric conversion is occurring faster than the NMR frequency. Furthermore, previous studies used a histidine dipeptide or tripeptide system, whereas native A $\beta$  consists of 40 residues with histidine residues at the 6, 13, and 14 positions. In this respect, previous NMR experiments may not reflect the true properties of native A $\beta$ , given that the properties of actual A $\beta$  are highly influenced by the electrostatic potential created by many side chains (17). This phenomenon implies that the IR or NMR signals of amide-I mode or the imidazole ring can be affected by the presence of other amino acids, especially when the histidine residue forms a secondary structure. At a neutral pH, histidine can exist as either a major  $\epsilon$  form or minor  $\delta$  form in a 1:0.16 ratio, whereas in acidic environments, histidine may be present in a protonated state (Fig. 1), although slight changes in the surrounding environment can perturb the ratio of histidine residues present. In addition, tautomeric equilibrium is reached on a picosecond timescale (18). Therefore, the current experimental approaches are somewhat difficult for providing a clear understanding of the tautomeric behavior of the histidine residues in A $\beta$ .

Femtosecond nonlinear two-dimensional infrared (2DIR) spectroscopy is a powerful tool for probing the dynamic properties of  $A\beta$  peptides in the amide-I frequency range. In particular, useful information regarding vibrational coupling between amide-I mode can be obtained by disentangling a congested one-dimensional spectrum into  $n$  dimensions to provide additional information on the secondary structures and their dynamics, such as inhomogeneous broadening and coupling between vibrational modes (19–21). Therefore, we performed histidine site-specific 2DIR calculations for amide-I bonds within the framework of classical MD simulation to investigate a possible connection between previous theoretical MD simulations and our experimental observations. We show that the  $\delta\delta\delta$  and  $\pi\pi\pi$  (all protonated) tautomers exhibit a greater propensity for misfolding compared to the  $\epsilon\epsilon\epsilon$  tautomer, as reported in our previous studies. We also report that the amide-I vibrational mode shows three distinctive peaks due to the tautomeric or protonation state of histidine side chains.

## METHODS

### Classical MD simulation

The initial tautomeric monomer structures were taken from our previous study (9), for which the crystal structure (PDB: 1BA4) was obtained from the Protein Data Bank. Herein, we refer to the protonated histidine residue as  $\pi\pi\pi$ , analogous to the notation for the  $\delta\delta\delta$  and  $\epsilon\epsilon\epsilon$  residues, although it is not the universally used name. The two tautomeric isomers,  $\delta\delta\delta$  and  $\epsilon\epsilon\epsilon$ , and the  $\pi\pi\pi$  protonated form (Fig. 1) were prepared by residual substitution with  $N^{\delta}$ -H or  $N^{\epsilon}$ -H at the 6, 13, and 14 histidine residues of  $A\beta$  (1–40). All MD simulations were performed with GROMACS 5.0 software (22) and the AMBER99SB force field (23). It was repeated with the CHARMM27 force field to validate our result. The monomer structures were solvated in TIP3P water with a minimal distance of 1.0 nm between structures and the edge of the box. Three  $Na^+$  ions were added to neutralize the system for the  $\delta\delta\delta$  and  $\epsilon\epsilon\epsilon$  tautomers. Long-range electrostatic interactions were treated through the particle mesh Ewald (PME) method with 0.16 nm spacing. Temperature was controlled by the V-rescale method (24). The SETTLE algorithm (25) was used to constrain water molecules. MD simulations were carried out for at least 200 ns with time step of 2 fs with the output of trajectory at every 20 ps. The last 200 ps trajectory with a 1 fs time step was used for further frequency and 2DIR calculations. The trajectories were analyzed using visual molecular dynamics (VMD) software (26). To analyze the secondary structure of each  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  tautomer, we used the DSSP method of Kabsch and Sander (27), which assigns secondary structure types for residues based on backbone amide (N-H) and carbonyl (C=O) positions.

To compare our results with previous experimental results (14), we further carried out classical MD simulation for histidine dipeptide monomers. The histidine dipeptide tautomers were named as  $HI\delta$  ( $\delta$  form),  $HI\epsilon$  ( $\epsilon$  form), and  $HI\pi$  (protonated form) tautomers depending on their protonated state of imidazole rings. All dipeptides were acetylated at the N-terminus and methyl amidated at the C-terminus like the experimental study (14). The same computational strategy was applied to histidine dipeptide monomers, including AMBER99SB force field, TIP3P water, PME method, and SETTLE algorithm, etc. MD simulations were performed for 200 ns with a time step of 2 fs, and the last 200 ps trajectory with a 1 fs time step was used for 2DIR calculations.

## Quantum calculation

Vibrational frequencies were calculated for  $HI\delta$ ,  $HI\epsilon$ , and  $HI\pi$  using MP2 theory with a basis set of 6-31G\*\* and a Gaussian package (28). A total of three types of models—histidine residue, histidine residue with implicit water, and the histidine dipeptide (same as classical MD model)—were considered. To distinguish two amide-I bonds of histidine dipeptide models, the amide-I bond located at the C-terminus was isotope labeled with  $C^{13}$  and  $O^{18}$ . The geometric structures were fully optimized, and anharmonic frequency calculations were carried out using the keyword `freq=anharmonic`, which uses numerical differentiation along modes, to determine the vibrational transitions of the amide-I bond.

## RESULTS AND DISCUSSION

Before analyzing the secondary structure of each tautomer, we calculated root mean-square deviations (RMSDs) to confirm equilibration of systems, using a backbone  $C_{\alpha}$  atom with respect to the initial structure for each  $\epsilon\epsilon\epsilon$ ,  $\delta\delta\delta$ , and  $\pi\pi\pi$  tautomer during the MD simulation (Fig. S1). The obtained RMSDs were less than 0.3 nm for all systems, indicating that our simulations had converged well. For  $\pi\pi\pi$ , we found an abrupt RMSD change near 200 ns; hence, we performed additional 200 ns MD simulations and only used the last 200 ns trajectories for analysis and frequency calculation.

Next, evolution of the secondary structures over time was investigated with the DSSP algorithm, as described in Methods. Formation of  $\beta$ -sheets is generally regarded as an important precursor of amyloid aggregation of  $A\beta$  (29–31). The average  $\beta$ -sheet,  $\alpha$ -helix, turn, and other characteristics were calculated as shown in Table 1. The  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  tautomers were composed of 17, 5, and 22%  $\beta$ -sheets, respectively. Specifically, the  $\delta\delta\delta$  tautomer contained  $\beta$ -sheet structures between the C-terminus (residue 4~5) and N-terminus (residue 37~38), as well as between the CHC (residue 16~21) and N-terminus (residue 32~37); in contrast, the  $\epsilon\epsilon\epsilon$  tautomer contained a  $\beta$ -bridge between residues 5~6 and 13~14, and the  $\pi\pi\pi$  tautomer contained a  $\beta$ -sheet between the CHC (residue 17~20) and N-terminus (residue 33~36) (Fig. 2). This trend has been confirmed by our previous studies, which used ff02 (9), AMBER99SBILDN (13), and ff14SB force fields (32). To validate our results, we further carried out the same MD simulation with the CHARMM27 force field (33). The CHARMM27 force field has been known to provide a quite different

**TABLE 1** Probability Distributions of Averaged Secondary Structures in the  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  Tautomers

	$\delta\delta\delta$	$\epsilon\epsilon\epsilon$	$\pi\pi\pi$
Coil	33%	31%	30%
$\beta$ -Sheet	17%	5%	22%
$\beta$ -Bridge	1%	6%	5%
Bend	11%	45%	17%
Turn	11%	3%	7%
$\alpha$ -Helix	16%	0%	12%
3-Helix	11%	9%	7%

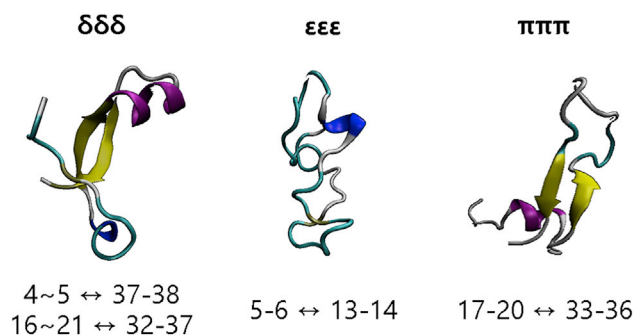


FIGURE 2 Representative conformational states of the  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  tautomers. The green band/line with arrow represents the  $\beta$ -sheet/bridge structure, and the numbers below indicate the residue numbers at which the  $\beta$ -sheet/bridge was formed. To see this figure in color, go online.

conformational state for amyloid  $\beta$  proteins (34). Our RMSD result confirms the equilibration of the systems (Fig. S1). The DSSP analysis showed that  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  exhibited 13, 0, and 16%  $\beta$ -sheet components, which is consistent with the simulation performed with the ff99SB force field. We expect that the conformational state can be varied by the use of different force fields: the use of the AMBER12SB force field could reduce the contents of  $\beta$ -sheet, whereas the use of AMBER96, GROMOS45a3, GROMOS53a5, GROMOS53a6, GROMOS43a1, GROMOS43a2, and GROMOS54a7 force field could increase the composition of  $\beta$ -sheet in light of the previous report (35) about the relation between force field and protein conformational state.

The contact map of each tautomer based on the  $C_\alpha$  atom (Fig. 3) supported the secondary structure conformations. The contact map for the  $\delta\delta\delta$  tautomer showed significant contacts between the 4~5 and 33~39 residues and between the 14~21 and 33~39 residues. The two regions in Fig. 3 marked with black circles show (anti)parallel conformations. In contrast, the  $\epsilon\epsilon\epsilon$  tautomer had some contacts between residues 4~7 and 12~15, indicating the presence of a  $\beta$ -bridge as illustrated in Fig. 2. This result is consistent with our previous theoretical calculations (9), indicating that  $\delta\delta\delta$  shows a greater propensity for  $\beta$ -sheet formation

compared with  $\epsilon\epsilon\epsilon$ . In addition, our results support previous experimental results indicating that  $\beta$ -sheets are easily produced between the CHC and C-terminus (36,37). Together, these findings indicate the influence of the tautomeric state of histidine on the structural properties and mechanism of A $\beta$  aggregation (8,32,38). The  $\pi\pi\pi$  tautomer was found in similar locations to and had secondary structures more similar to those found in the  $\delta\delta\delta$  tautomer than the  $\epsilon\epsilon\epsilon$  tautomer. Comparison of the  $\pi\pi\pi$  system with the tautomeric monomers ( $\delta\delta\delta$  and  $\epsilon\epsilon\epsilon$ ) has not been reported. On pathological examination, the brains of AD patients are relatively acidic compared with the brains of normal people. In an acidic environment, the histidine residue may be protonated with an  $N^\delta$ -H bond. Comparing the pathological observations and the similarity of MD simulations for the  $\pi\pi\pi$  and  $\delta\delta\delta$  tautomers, the presence of  $N^\delta$ -H might lead to greater formation of  $\beta$ -sheet structures. Hence, it is important to distinguish the two tautomeric  $N^\delta$ -H and  $N^\epsilon$ -H forms using spectroscopic methods.

In addition to the contact maps, we calculated the strength of dipole coupling for the amide-I vibrational modes. We used an equivalent magnitude of transition dipole moment for all peptide bonds so that the coupling strength was attributed to the related secondary structure. The strength of the transition dipole was assumed to be 0.374 D, with the direction being indicated from the C=O bond toward the C-N bond at the same plane with a 20.0° angle (Fig. S3). This assumption was first suggested by Krimm et al. (39); since then, the strength of transition dipole and its angle against the C=O bond have been modified by many authors (40–42). It has been known that this approach provides qualitatively similar results compared with other method such as dipole approximation, transition charge coupling, local (normal) mode molecular orbital calculation, etc. (42–45). The optimized location of the transition dipole was assumed to be  $r_C + 0.665n_{CO} + 0.258n_{CN}$ , where  $r_C$  represents the location of the carbon atom of the amide-I bond,  $n_{CO} = (r_O - r_C)/|r_O - r_C|$ , and  $n_{CN} = (r_N - r_C)/|r_N - r_C|$  (Fig. S3). The strength of dipole coupling ( $\beta_{jk}$ ) between two amide-I modes was then calculated as

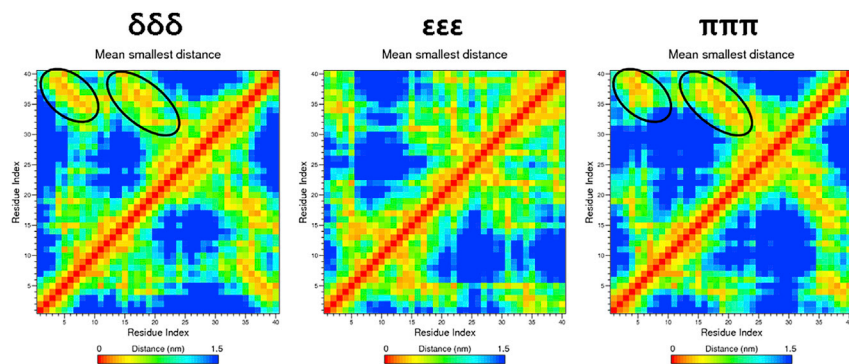


FIGURE 3 Contact maps for the  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  tautomers. The  $x$  and  $y$  axes represent the residue numbers for each tautomer. Black circles represent the locations of the antiparallel  $\beta$ -sheets/bridges shown in Fig. 2. To see this figure in color, go online.



$$\beta_{jk} = \frac{1}{4\pi\epsilon_0} \times \frac{(\partial\mu_j/\partial Q_j)(\partial\mu_k/\partial Q_k) - 3[(\partial\mu_j/\partial Q_j)n_{jk}][(\partial\mu_k/\partial Q_k)n_{jk}]}{R_{jk}^3}, \quad (1)$$

where  $(\partial\mu_\alpha/\partial Q_\alpha)$  ( $\alpha = j, k$ ) is the dipole derivative (transition dipole in  $\text{D } \text{\AA}^{-1} \text{ u}^{-1/2}$ ) of the amide-I vibration of the  $\alpha^{\text{th}}$  peptide group,  $R_{jk}$  is the length (in  $\text{\AA}$ ) of the line connecting the  $j^{\text{th}}$  and  $k^{\text{th}}$  peptide groups,  $n_{jk}$  is the unit vector along this line, and  $\epsilon_0$  is the vacuum permittivity. The strength of dipole coupling was calculated for all pairs of amide groups at every trajectory and then averaged to obtain the coupling map as shown in Fig. 4. The coupling maps of the  $\delta\delta\delta$  and  $\pi\pi\pi$  tautomers clearly showed strong dipole couplings between the N-terminus and C-terminus, as well as between the CHC and C-terminus, as marked with red circles in the figure. The dipole coupling at circled area showed  $-10$  to  $-15 \text{ cm}^{-1}$ , which corresponds to the general dipole coupling strength between antiparallel  $\beta$ -sheets, which ranges from  $-10$  to  $-20 \text{ cm}^{-1}$  (46,47). In addition, the distances between transition dipoles ranged from 0.45 to 0.50 nm (Table S1), similar to the distance between antiparallel  $\beta$ -sheets of 0.47–0.48 nm (48). However, the coupling map of the  $\epsilon\epsilon\epsilon$  tautomer does not exhibit such strong coupling in that region. The (anti)parallel  $\beta$ -sheet yields a stronger oscillator coupling compared with other secondary structures because the  $\beta$ -strands are arranged adjacent to other strands and form an extensive hydrogen bond network with their neighbors. By comparing this result with Fig. 2, our transition dipole coupling map results provide a clue for the location of antiparallel  $\beta$ -sheet.

The above findings, which support the tautomer hypothesis, have been supported by our previous studies (8,9,12,32,38) addressing homodimer and pentamer aggregation mechanisms using MD simulations. Experimentally, it has been reported that  $\text{N}^\delta\text{-H}$  and  $\text{N}^\epsilon\text{-H}$  of the histidine dipeptide and tripeptide can be distinguished using NMR techniques; linear IR and 2DIR (14) techniques have also been used to probe the tautomeric form of the histidine residues of the dipeptide. However, the previous results may not accurately reflect the actual  $\text{A}\beta$  system composed of

40–42 residues including three histidine residues. In addition, the  $\text{A}\beta$  structure is affected by surrounding water molecules on a femtosecond scale. Therefore, we calculated the frequencies and 2DIR spectra of the  $\text{A}\beta$  structure upon histidine tautomeric state to provide an experimentally comparable result in the following section.

## Frequency calculations

Several parameterization schemes have been used to describe the calculation of the amide-I vibrational frequencies. Among them, we used CHO4 parametrization (49–51). The amide-I vibrational frequency ( $\hbar\omega_n$ ) of the  $n^{\text{th}}$  peptide bond located between residue  $r$  and  $r + 1$  was calculated as

$$\hbar\omega_n = \hbar\omega_0 + \sum_{s=C,O,N,H} l_s \phi_{n,s}(t). \quad (2)$$

The sum is performed over the C, O, N, and H atoms of the amide bond  $n$ . The central band ( $= \hbar\omega_0$ ) frequency was set to  $1600 \text{ cm}^{-1}$  with reference to the previous simulation study (17), which performed frequency calculation of the same  $\text{A}\beta$  (1–40) system. The values of the coefficient  $l_s$  are given as  $l_C = -0.00554e$ ,  $l_O = 0.00160e$ ,  $l_N = 0.00479e$ , and  $l_H = -0.00086e$ , where  $e$  is the electron charge. The electrostatic potential  $\phi_{n,s}$  at the  $r_{n,s}(t)$  coordinate of atom  $s$  of the  $n^{\text{th}}$  amide bond was expressed as the sum of the contributions from water, backbone, and side chains, as follows:

$$\phi_{n,s} = \phi_{n,s}^{\text{water}}(t) + \phi_{n,s}^{\text{side chains}}(t) + \phi_{n,s}^{\text{backbone}}(t), \quad (3)$$

$$\phi_{n,s}^P(t) = \frac{1}{4\pi\epsilon_0} \sum_{i \in P} \frac{q_i}{|r_i(t) - r_{n,s}(t)|}, \quad (4)$$

( $P = \text{Water, Side chains, or backbone}$ ),

where  $\phi_{n,s}^{\text{water}}(t)$ ,  $\phi_{n,s}^{\text{side chains}}(t)$ , and  $\phi_{n,s}^{\text{backbone}}(t)$  represent the electrostatic potential generated by the surrounding water molecules, side chains, and backbone atoms, respectively. The electrostatic potential contribution is calculated by sum of point charges,  $q_i$ , over distance  $|r_i(t) - r_{n,s}(t)|$  and

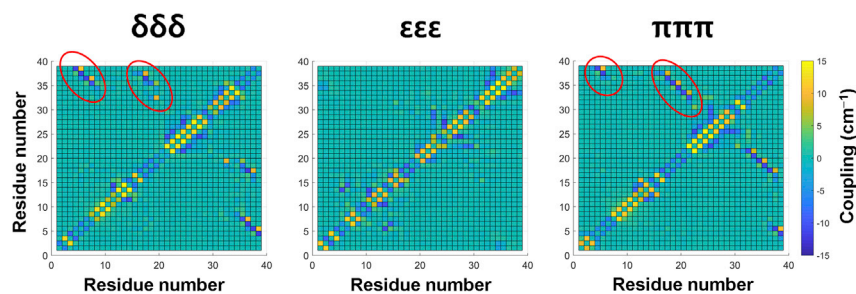


FIGURE 4 Coupling maps showing the strength of oscillator couplings between amide-I bonds (unit:  $\text{cm}^{-1}$ ). The  $x$  and  $y$  axes represent the residue numbers of each tautomer. Red circles represent the location of the antiparallel  $\beta$ -sheets/bridges shown in Fig. 2. To see this figure in color, go online.

the vacuum permittivity  $\epsilon_0$  for all atoms ( $i$ ) in the water, side chain, or backbone atoms. The  $q_i$  of water molecules were calculated using TIP3P point charge, and that of the side chain was given by the ff99SB force field. Finally, the electrostatic potential generated by the amyloid backbone atoms, C, O, N, H,  $C_\alpha$ , and  $H_\alpha$ , was calculated based on Ham's model (49,50), in which no effective charges are assigned to  $C_\alpha$  or  $H_\alpha$  atoms ( $q_{C_\alpha} = 0, q_{H_\alpha} = 0$ ), and  $q_C = 0.419, q_O = -0.871, q_N = 0.793$ , and  $q_H = -0.341$ .

The average frequency shift,  $\langle \hbar\omega_0 - \hbar\omega_n \rangle$ , of amide-I vibration induced by the electrostatic potential of the backbone, side chain, and water is shown in Fig. 5. We first noticed a large shift in both the N- and C-terminus, which corresponded to the head and tail of tautomers, respectively. Further, the frequency shift in the regions between residues Asp1–Arg5 was larger than for other regions, indicating the disordered structure of the N-terminus, as reported in previous studies (52,53). The  $\sim -20 \text{ cm}^{-1}$  frequency shift for the backbone was similar for all tautomers. This finding is consistent with the previous observation that most of the inhomogeneity of the frequency shift is dominated by the side chain contribution (17) and not by the backbone contribution, which further supports the importance of tautomeric form of histidine side chains for A $\beta$  misfolding and aggregation. Next, we investigated the histidine site-specific 2DIR spectrum of each tautomer; however, for these analyses, we disregarded the water contribution to assess dried A $\beta$  samples containing only one or two water molecules (54).

## 2DIR spectra

The 2DIR spectra were calculated using snapshots containing the coordinates of the  $\epsilon\epsilon\epsilon$ ,  $\delta\delta\delta$ , and  $\pi\pi\pi$  tautomers obtained at every 1 fs time step during the final 200 ps of their frequency trajectories. The amide-I vibrational anharmonicity was fixed at  $-16 \text{ cm}^{-1}$  as a typical value for peptides. The frequency fluctuation  $\delta\omega(t) = \omega(t) - \langle \omega(t) \rangle$  was obtained from the frequency trajectories  $\omega(t)$  calculated in the previous section, where  $\langle \omega(t) \rangle$  was the ensemble average. The frequency fluctuation correlation function

(FFCF)  $\langle \delta\omega(0)\delta\omega(t) \rangle$  was then calculated and used to obtain the lineshape function (21),  $g(t)$ , as

$$g(t) = \int_0^t d\tau' \int_0^{\tau'} d\tau'' \langle \delta\omega(0)\delta\omega(t) \rangle, \quad (5)$$

from which the linear response function and the rephasing and non-rephasing third-order response functions were obtained using a cumulant approximation (21), as follows:

$$\begin{aligned} R_{1,2,3} &\propto i\mu_{01}^4 \left( e^{-i\omega_{01}(t_3-t_1)} - e^{-i((\omega_{01}-\Delta)t_3-\omega_{01}t_1)} \right) \\ &\times e^{-g(t_1)+g(t_2)-g(t_3)-g(t_1+t_2)-g(t_2+t_3)+g(t_1+t_2+t_3)}, \\ R_{4,5,6} &\propto i\mu_{01}^4 \left( e^{-i\omega_{01}(t_3+t_1)} - e^{-i((\omega_{01}-\Delta)t_3+\omega_{01}t_1)} \right) \\ &\times e^{-g(t_1)-g(t_2)-g(t_3)+g(t_1+t_2)+g(t_2+t_3)-g(t_1+t_2+t_3)}, \end{aligned} \quad (6)$$

where  $\Delta$ ,  $\mu_{01}$ , and  $\omega_{01}$  refer to vibrational anharmonicity ( $-16 \text{ cm}^{-1}$ ), the time-independent vibrational transition dipole moment, and the time-independent mean frequency, respectively. The FFCF can be integrated into the lineshape function either by numerically or analytically fitted to a parametrized form:

$$c(t) = \langle \delta\omega(0)\delta\omega(t) \rangle = a_1 \cos(\gamma t)e^{-t/\tau_1} + a_2 e^{-t/\tau_2} + a_3 e^{-t/\tau_3}, \quad (7)$$

where  $a_1, a_2, a_3, \tau_1, \tau_2, \tau_3$ , and  $\gamma$  are parameters (21,55).

During the numerical integration, the direct relationship between the signal and FFCF is lost (56). For example, the first term of the correlation function is an attempt to catch the shorter time dynamics related to hydrogen bonded species, such as solvent fluctuations (57,58). The remaining two terms also provide a slower dynamic property of the system. Usually, an analytical integration cannot be obtained because the lineshape function shows a complicated set of integrals of exponential functions, but our FFCFs were well fitted to the parametrized correlation function;

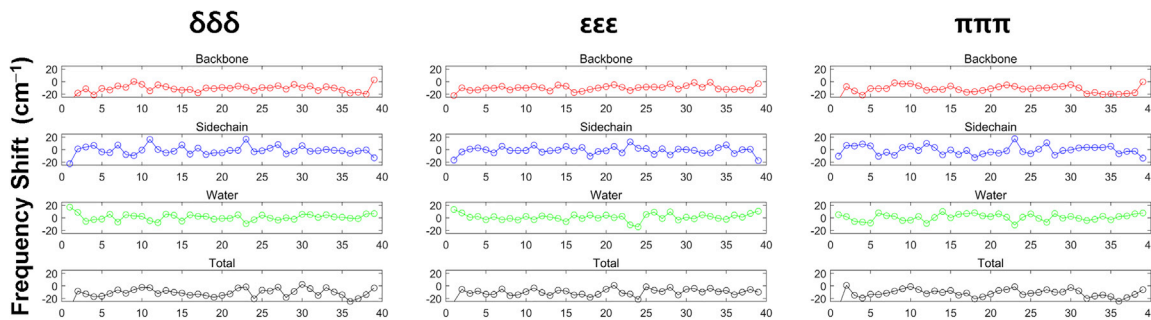


FIGURE 5 Average frequency shifts of the  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  tautomers due to the backbone (red line), side chains (blue line), water solvent (green line), and entire system (black line). To see this figure in color, go online.

hence, the analytically integrated lineshape function was used to calculate response function as described above. Once the response functions  $R_1$  to  $R_6$  are known, the frequency-domain spectra are calculated via Fourier transformation of response function over  $t_1$  and  $t_3$ .

$$S_{r/nr}(\omega_1, t_2, \omega_3) = \int_0^\infty \int_0^\infty R_{r/nr}(t_1, t_2, t_3) e^{i\omega_1 t_1} e^{i\omega_3 t_3} dt_1 dt_3, \quad (8)$$

where r/nr refers to rephasing and non-rephasing, respectively. After inverting the sign of  $\omega_1$  of the rephasing response function, the real part ( $\mathbb{R}$ ) of the sum of both rephasing and non-rephasing spectra is taken to calculate the purely absorptive 2DIR spectra,

$$S_{abs}(\omega_1, t_2, \omega_3) = \mathbb{R}[S_r(-\omega_1, t_2, \omega_3) + S_{nr}(\omega_1, t_2, \omega_3)]. \quad (9)$$

The spectra were obtained for waiting time  $t_2 = 0$  ps in this study.

The site-specific 2DIR spectra at His6, His13, and His14 of each tautomer and their diagonal traces of 2DIR spectra at zero waiting time,  $t_2 = 0$ , are shown in Fig. 6. The diagonal traces for all the tautomers contain bands at  $\sim 1580$ ,  $\sim 1586$ , and  $1593 \text{ cm}^{-1}$ , similar to previously reported experimental results (14) ( $1630$ ,  $1644$ , and  $1656 \text{ cm}^{-1}$ ). The smaller peaks of our system compared with experimental data originated from isotope labeling of the amide-I bond, which reduces the vibrational band by  $60 \text{ cm}^{-1}$ . This isotope frequency shift of  $60 \text{ cm}^{-1}$  was reported by previous studies (21,59) and provides the advantage of placing the isotope-labeled amide-I band between the unlabeled amide-I and II bands. The calculated bands can vary with the use of the different isotope frequency shift or the central band frequency,  $\hbar\omega_0$ ; however, our results provide three resolvable amide-I spectra upon tautomerization or protonation, which validates our results. Specifically, at His6, the  $\epsilon\epsilon\epsilon$  tautomer exhibited a slightly higher peak ( $1586 \text{ cm}^{-1}$ ) than the  $\delta\delta\delta$  and  $\pi\pi\pi$  tautomers ( $1580 \text{ cm}^{-1}$ ). In contrast, the  $\epsilon\epsilon\epsilon$  tautomer had a lower peak ( $1580 \text{ cm}^{-1}$ ) than the  $\delta\delta\delta$  ( $1586 \text{ cm}^{-1}$ ) and  $\pi\pi\pi$  ( $1593 \text{ cm}^{-1}$ ) tautomers at His13. Again,  $1593 \text{ cm}^{-1}$  and His14 peaks for the  $\epsilon\epsilon\epsilon$  tautomer were higher than those for the  $\delta\delta\delta$  ( $1586 \text{ cm}^{-1}$ ) and  $\pi\pi\pi$  ( $1580 \text{ cm}^{-1}$ ) tautomers. These differences in diagonal peaks for the tautomeric/protonated state of the histidine residues of A $\beta$  (1–40) have not previously been reported and are entirely new findings, to our knowledge.

According to the previous experimental report (14) the HI $\delta$  dipeptide (histidine with N-H $\delta$  bond) has a higher ( $\sim 1656 \text{ cm}^{-1}$ ) peak compared with that of the HI $\epsilon$  dipeptide (N-H $\epsilon$  bond). A density functional theory (DFT) study indicates the HI $\delta$  dipeptide exhibits an amide-I peak  $\sim 5 \text{ cm}^{-1}$  higher than that of the HI $\epsilon$  dipeptide. To verify our results,

we also performed ab initio electronic structure and frequency calculations (see Methods) and summarized the results in Table S2. The HI $\delta$  and HI $\pi$  tautomers exhibited higher amide-I peaks compared to the HI $\epsilon$  tautomer ( $\sim 10$  and  $12 \text{ cm}^{-1}$ , respectively). However, when the solvent effect was included using an implicit water solvent model, the peak for the HI $\pi$  tautomer was similar to that for the HI $\epsilon$  tautomer, whereas the HI $\delta$  tautomer peak remained  $\sim 9 \text{ cm}^{-1}$  higher. In contrast, the dipeptide system, which included two amide bonds, showed the opposite trend: the two amide bonds in the dipeptide system were distinguishable by peak position because the vibrational frequency of the isotope-labeled amide-I bond (amidated C-terminus) was observed at the lower peak position compared with that of the nonlabeled bond (acetylated N-terminus). The relative magnitudes of the peak positions of the isotope-labeled amide-I bond were HI $\epsilon >$  HI $\delta >$  HI $\pi$ , whereas the peaks for the nonlabeled amide-I bond were HI $\delta >$  HI $\epsilon >$  HI $\pi$ .

Next, we performed classical MD simulations for the HI $\delta$ , HI $\epsilon$ , and HI $\pi$  tautomers using the same methodology applied to the  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  A $\beta$  (1–40) systems (see Methods). The RMSDs obtained were less than  $0.1 \text{ nm}$  for all histidine dipeptides, indicating that our simulations had converged well (Fig. S2). Then, we calculated the 2DIR spectra and their diagonal traces of the HI $\delta$ , HI $\epsilon$ , and HI $\pi$  histidine dipeptide systems as shown in Fig. 7. Herein, we name the two amide-I bonds as the N-terminus amide-I bond and the C-terminus amide-I bond. The bottom left figure shows the diagonal trace of the C-terminus amide-I bond, and the right figure shows the diagonal trace of the N-terminus amide-I bond. We found that the diagonal peak for histidine can be characterized by three peaks at  $\sim 1580$ ,  $\sim 1586$ , and  $1593 \text{ cm}^{-1}$ . Specifically, the C-terminus amide-I bond for HI $\epsilon$  has a peak at  $1586 \text{ cm}^{-1}$ , whereas HI $\delta$  and HI $\pi$  have peaks at  $1580 \text{ cm}^{-1}$ . In contrast, the  $1600 \text{ cm}^{-1}$  peak for the N-terminus amide-I bond in HI $\epsilon$  was lower than those for HI $\delta$  or HI $\pi$  ( $1607$  or  $1613 \text{ cm}^{-1}$ , respectively).

The above results indicated that the peak position for each tautomer may be highly dependent on the environment, such as the surrounding hydrogen bonds. The presence of a hydrogen bond lowers the peak position because they weaken the C=O bond strength. In this context, we calculated the average number of hydrogen bonds, for which the distance between donor (OH or NH groups) and acceptor (O, N atoms) is smaller than  $0.35 \text{ nm}$ , during the dynamics of each HI $\delta$ , HI $\epsilon$ , and HI $\pi$  system (Table S3). The result shows that the average number of hydrogen bonds for the C-terminus amide bond is smaller for HI $\epsilon$  (2.054) than for HI $\delta$  (2.967) or HI $\pi$  (2.533), whereas that for the N-terminus-amide bond is larger for HI $\epsilon$  (3.450) than HI $\delta$  (3.413) or HI $\pi$  (3.109). This trend was consistent with the diagonal peak positions of the HI $\delta$ , HI $\epsilon$ , and HI $\pi$  tautomer systems shown in Fig. 7. The protonation of the histidine imidazole

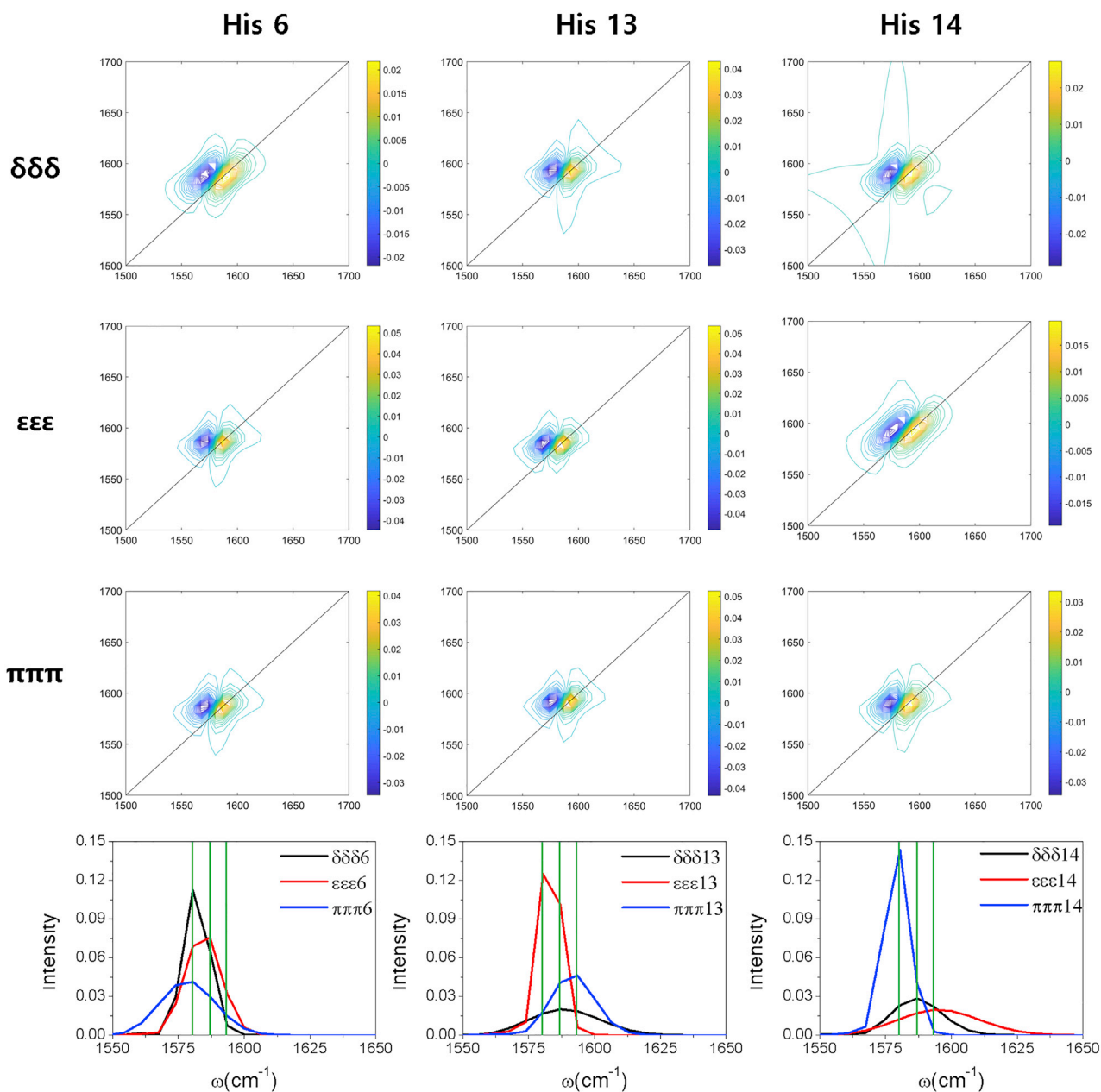


FIGURE 6 Histidine site-specific 2DIR spectra and their diagonal traces of the  $\delta\delta\delta$ ,  $\varepsilon\varepsilon\varepsilon$ , and  $\pi\pi\pi$  tautomers for the 6th, 13th, and 14th amide-I bond of histidine residue. The black, red, and blue lines represent the diagonal traces of the  $\delta\delta\delta$ ,  $\varepsilon\varepsilon\varepsilon$ , and  $\pi\pi\pi$  tautomers, respectively. The green lines represent the three peaks at 1580, 1586, and 1593  $\text{cm}^{-1}$ . To see this figure in color, go online.

ring results in charge redistribution in histidine amino acid. As a result, the  $\text{HI}\pi$  tautomer has more positive C atoms and more negative O atoms compared with those of  $\text{HI}\delta$  and  $\text{HI}\varepsilon$  (Table S4). This indicates that for the C-terminus amide-I bond, the  $\text{HI}\pi$  tautomer shows larger amide-I frequency because of a strong C=O bond than  $\text{HI}\delta$  or  $\text{HI}\varepsilon$ , which might lead to the higher propensity of hydrogen bond formation because of a strongly polarized O atom and its frequent oscillation. However, the larger number of hydrogen bonds for the  $\text{HI}\pi$  former amide bond than  $\text{HI}\varepsilon$  shifts the peak po-

sition lower than  $\text{HI}\varepsilon$ , resulting in the higher peak position for  $\text{HI}\varepsilon$  than  $\text{HI}\delta$  and  $\text{HI}\pi$ . In contrast, for the N-terminus amide-I bond, all three tautomers intrinsically have the same amide-I frequency, but the larger number of hydrogen bonds lowers the peak position of  $\text{HI}\varepsilon$  more than  $\text{HI}\delta$  and  $\text{HI}\pi$ .

However, there are several reasons that we cannot directly compare our result with experimental observations. First, site-specific isotope labeling of histidine residue in polypeptides or proteins has not been reported yet. Ghosh's work



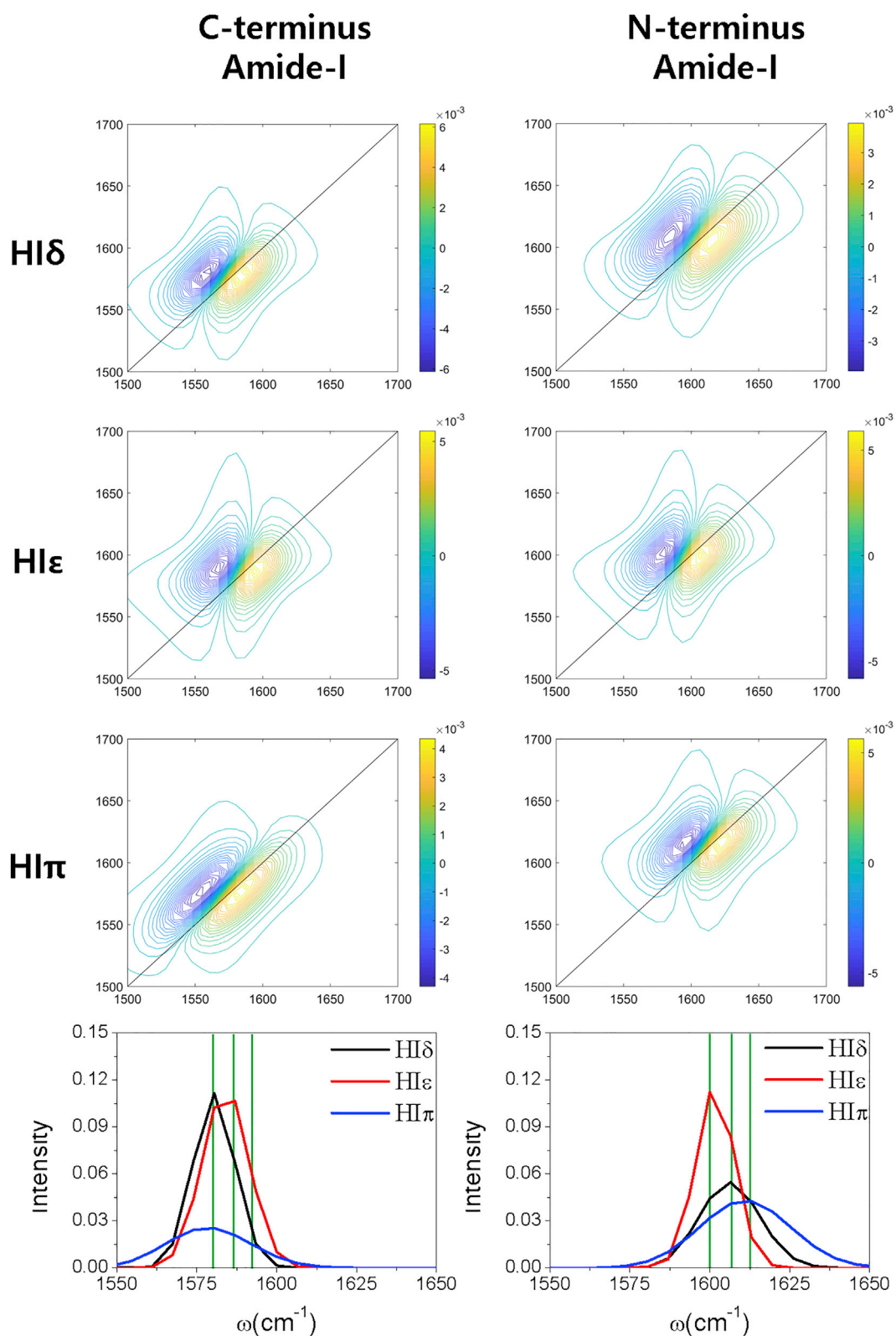


FIGURE 7 2DIR spectra and their diagonal traces of 2DIR spectra for amide-I vibration of  $\text{HI}\delta$ ,  $\text{HI}\epsilon$ , and  $\text{HI}\pi$  dipeptides: C-terminus (*left*) and N-terminus (*right*) amide bonds. Black, red, and blue lines represent the diagonal traces of the  $\text{HI}\delta$ ,  $\text{HI}\epsilon$ , and  $\text{HI}\pi$  dipeptides, respectively. Green lines represent the three peaks at 1580, 1586, and 1593  $\text{cm}^{-1}$  for isotope-labeled bonds and 1600, 1607, and 1613  $\text{cm}^{-1}$  for nonlabeled amide bonds. To see this figure in color, go online.

(14) is the only existing experimental study that compares IR and 2DIR spectra upon tautomerization or protonation. Hence, the comparison of our result is somewhat limited. Second, the 2DIR spectra of Ghosh's work was obtained at different pH (e.g., pH 10, 6.5, and 2), whereas our system was neutralized. Third, the amide vibration of a histidine residue is an indirect probe of the side chain's conformational state. Hence, Ghosh's work (14) performed quantum calculation to calculate the amide-I vibrational frequency, but their DFT results were different from our results, which were performed based on wavefunction electronic structure method. Even the vibrational frequency within MP2 calculation was highly dependent on the solvation state. Finally, we used classical MD simulation to calculate 2DIR spectra of histidine dipeptide, which cannot elucidate the chemical exchanges. Hence, our 2DIR spectra at waiting time,  $t_2 = 3$  ps, did not show crosspeaks as observed by experiments (Fig. S4). Nevertheless, we showed that our histidine dipeptide exhibits three spectrally resolvable features within the amide-I vibrational band of histidine.

We believe that the resolvable three peaks could be more accurately elucidated by using *ab initio* MD simulation or density functional tight-binding MD simulation, which can elucidate the bond-breaking phenomenon (intermolecular proton transfer process). The amide-I vibrational peak is highly influenced by the environment: our  $\delta\delta\delta/\epsilon\epsilon\epsilon/\pi\pi\pi$  and  $\text{HI}\delta/\text{HI}\epsilon/\text{HI}\pi$  systems had different geometric structures, leading to different dynamics scenarios. However, the change of amide-I peak position, as well as the timescale of proton transfer, could be elucidated by investigating the 2DIR spectra of each system, which will be performed on a future study. We expect that this study could inspire further studies on the distinction of amide-I vibration by protonation or tautomerization of histidine side chains.

## CONCLUSION

In summary, we calculated histidine site-specific 2DIR spectra within a classical MD simulation framework to investigate the effect of histidine tautomeric state on amyloid  $\beta$ -sheet misfolding and aggregation of A $\beta$  (1–40). We found that the  $\delta\delta\delta$  (17%) and  $\pi\pi\pi$  (22%) tautomers exhibited greater propensity for  $\beta$ -sheet formation between the C-terminus and N-terminus, as well as between the CHC and N-terminus, as compared with the  $\epsilon\epsilon\epsilon$  tautomer (5%). In particular, investigations of the dominant secondary structure for the  $\pi\pi\pi$  tautomer have not been previously reported. This trend was further confirmed by contact map analyses and dipole coupling mapping, which clearly showed the presence of (anti)parallel  $\beta$ -sheets for the  $\delta\delta\delta$  and  $\pi\pi\pi$  tautomers. The histidine site-specific 2DIR diagonal peak of each tautomer showed three distinctive peaks near 1580, 1586, and 1593  $\text{cm}^{-1}$ , as observed in the experiments for histidine dipeptides. However, the diagonal peaks did not show a similar trend for the histidine or protonation

state and were instead affected by the presence of surrounding hydrogen bonds. In the future, obtaining a more detailed understanding of tautomerism for histidine residues may be possible as a result of performing *ab initio* MD simulations, which can elucidate proton exchange between histidine side chains and surrounding water molecules. Nevertheless, our study is the first, to our knowledge, to investigate the histidine site-specific 2DIR spectra for the amide-I vibrational mode of A $\beta$  (1–40) in the histidine tautomeric state. We expect these results may validate 2DIR spectroscopy as a possible tool for studying the influence of aggregation of A $\beta$  (1–40) arising from histidine tautomerism.

## SUPPORTING MATERIAL

Supporting Material can be found online at <https://doi.org/10.1016/j.bpj.2020.07.009>.

## AUTHOR CONTRIBUTIONS

Y.N. and J.Y.L. performed research, analyzed data, and wrote the manuscript. M.K. contributed to data analysis by analytic tools. S.S. and J.Y.L. designed research and wrote the manuscript.

## ACKNOWLEDGMENTS

Y.N. acknowledges Toshifumi Mori for his assistance with the use of the supercomputers and computer environment.

This work was supported by Korea Initiative for fostering University of Research and Innovation Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. 2020M3H1A1077095) and National Research Foundation (NRF) grants funded by the Korean government (NRF-2019R1A6A1A10073079). S.S. was supported by JSPS KAKENHI Grant No. JP16H02254. The authors would like to acknowledge support from KISTI Supercomputing Center through the strategic support program for supercomputing application research (No. KSC-2018-CHA-0025) and also the support from the Research Center for Computational Science in Okazaki for calculation of the 2DIR spectra.

## REFERENCES

- Alonso, A., T. Zaidi, ..., K. Iqbal. 2001. Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments. *Proc. Natl. Acad. Sci. USA*. 98:6923–6928.
- Hardy, J., and D. J. Selkoe. 2002. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*. 297:353–356.
- Zhang, C.-E., W. Wei, ..., J.-Z. Wang. 2009. Hyperhomocysteinemia increases beta-amyloid by enhancing expression of gamma-secretase and phosphorylation of amyloid precursor protein in rat brain. *Am. J. Pathol.* 174:1481–1491.
- Shi, H., B. Kang, and J. Y. Lee. 2014. Zn(2+) effect on structure and residual hydrophobicity of amyloid  $\beta$ -peptide monomers. *J. Phys. Chem. B*. 118:10355–10361.
- Zhang, R., H. Ai, ..., Q. Li. 2016. Molecular simulations of human and mouse A $\beta$ 1–16 at different pH values: structural characteristics toward understanding Cu(2+) -coordinated amyloid beta spheres. *ChemPhysChem*. 17:1656–1668.

6. Burdick, D., B. Soreghan, ..., C. Glabe. 1992. Assembly and aggregation properties of synthetic Alzheimer's A4/beta amyloid peptide analogs. *J. Biol. Chem.* 267:546–554.
7. Yates, C. M., J. Butterworth, ..., A. Gordon. 1990. Enzyme activities in relation to pH and lactate in postmortem brain in Alzheimer-type and other dementias. *J. Neurochem.* 55:1624–1630.
8. Shi, H., and J. Y. Lee. 2017. Tautomeric effect of histidine on the monomeric structure of amyloid  $\beta$ -peptide(1-42). *ACS Chem. Neurosci.* 8:669–675.
9. Shi, H., B. Kang, and J. Y. Lee. 2016. Tautomeric effect of histidine on the monomeric structure of amyloid  $\beta$ -peptide(1-40). *J. Phys. Chem. B.* 120:11405–11411.
10. Hansen, A. L., and L. E. Kay. 2014. Measurement of histidine pKa values and tautomer populations in invisible protein states. *Proc. Natl. Acad. Sci. USA.* 111:E1705–E1712.
11. Olubiyi, O. O., and B. Strodel. 2012. Structures of the amyloid  $\beta$ -peptides A $\beta$ 1-40 and A $\beta$ 1-42 as influenced by pH and a D-peptide. *J. Phys. Chem. B.* 116:3280–3291.
12. Salimi, A., H. Li, ..., J. Y. Lee. 2019. Intrinsic origin of amyloid aggregation: behavior of histidine ( $\epsilon\epsilon\epsilon$ ) and ( $\delta\delta\delta$ ) tautomer homodimers of A $\beta$  (1–40). *Biochim. Biophys. Acta Gen. Subj.* 1863:795–801.
13. Xing, X., W. Zhao, ..., H. Ai. 2019. Tautomerization effect of histidines on oligomer aggregation of  $\beta$ -amyloid(1-40/42) during the early stage: tautomerism hypothesis for misfolding protein aggregation. *ACS Chem. Neurosci.* 10:2602–2608.
14. Ghosh, A., M. J. Tucker, and F. Gai. 2014. 2D IR spectroscopy of histidine: probing side-chain structure and dynamics via backbone amide vibrations. *J. Phys. Chem. B.* 118:7799–7805.
15. Cheng, F., H. Sun, ..., E. Oldfield. 2005. A solid state  $^{13}\text{C}$  NMR, crystallographic, and quantum chemical investigation of chemical shifts and hydrogen bonding in histidine dipeptides. *J. Am. Chem. Soc.* 127:12544–12554.
16. Vila, J. A., Y. A. Arnautova, ..., H. A. Scheraga. 2011. Assessing the fractions of tautomeric forms of the imidazole ring of histidine in proteins as a function of pH. *Proc. Natl. Acad. Sci. USA.* 108:5602–5607.
17. Falvo, C., W. Zhuang, ..., S. Mukamel. 2012. Frequency distribution of the amide-I vibration sorted by residues in amyloid fibrils revealed by 2D-IR measurements and simulations. *J. Phys. Chem. B.* 116:3322–3330.
18. Nydegger, M. W., S. Dutta, and C. M. Cheatum. 2010. Two-dimensional infrared study of 3-azidopyridine as a potential spectroscopic reporter of protonation state. *J. Chem. Phys.* 133:134506.
19. Jansen, T. C., S. Saito, ..., M. Cho. 2019. Theory of coherent two-dimensional vibrational spectroscopy. *J. Chem. Phys.* 150:100901.
20. Yagasaki, T., and S. Saito. 2013. Fluctuations and relaxation dynamics of liquid water revealed by linear and nonlinear spectroscopy. *Annu. Rev. Phys. Chem.* 64:55–75.
21. Hamm, P., and M. Zanni. 2011. Concepts and Methods of 2D Infrared Spectroscopy. Cambridge University Press, Cambridge, UK.
22. Abraham, M. J., T. Murtola, ..., E. Lindahl. 2015. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 1–2:19–25.
23. Hornak, V., R. Abel, ..., C. Simmerling. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins.* 65:712–725.
24. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101.
25. Miyamoto, S., and P. A. Kollman. 1992. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 13:952–962.
26. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.
27. Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.
28. Frisch, M. J., G. W. Trucks, ..., D. J. Fox. 2016. Gaussian 16, Revision C.01. Gaussian Inc., Wallingford, CT.
29. Redler, R. L., D. Shirvanyants, ..., N. V. Dokholyan. 2014. Computational approaches to understanding protein aggregation in neurodegeneration. *J. Mol. Cell Biol.* 6:104–115.
30. David, D. C., N. Ollikainen, ..., C. Kenyon. 2010. Widespread protein aggregation as an inherent part of aging in *C. elegans*. *PLoS Biol.* 8:e1000450.
31. Liao, M. Q., Y. J. Tzeng, ..., Y. C. Chen. 2007. The correlation between neurotoxicity, aggregative ability and secondary structure studied by sequence truncated Abeta peptides. *FEBS Lett.* 581:1161–1165.
32. Shi, H., H. Li, ..., W. Guo. 2019. Structural and binding properties on A $\beta$  mature fibrils due to the histidine tautomeric effect. *ACS Chem. Neurosci.* 10:4612–4618.
33. MacKerell, A. D., D. Bashford, ..., M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
34. Weber, O. C., and V. N. Uversky. 2017. How accurate are your simulations? Effects of confined aqueous volume and AMBER FF99SB and CHARMM22/CMAP force field parameters on structural ensembles of intrinsically disordered proteins: amyloid- $\beta_{42}$  in water. *Intrinsically Disord. Proteins.* 5:e1377813.
35. Man, V. H., X. He, ..., J. Wang. 2019. Effects of all-atom molecular mechanics force fields on amyloid peptide assembly: the case of A $\beta_{16-22}$  dimer. *J. Chem. Theory Comput.* 15:1440–1452.
36. Sgourakis, N. G., Y. Yan, ..., A. E. Garcia. 2007. The Alzheimer's peptides Abeta40 and 42 adopt distinct conformations in water: a combined MD / NMR study. *J. Mol. Biol.* 368:1448–1457.
37. Sgourakis, N. G., M. Merced-Serrano, ..., A. E. Garcia. 2011. Atomic-level characterization of the ensemble of the A $\beta$ (1–42) monomer in water using unbiased molecular dynamics simulations and spectral algorithms. *J. Mol. Biol.* 405:570–583.
38. Li, H., A. Salimi, and J. Y. Lee. 2019. Intrinsic origin of amyloid aggregation: collective effects of the mutation and tautomerism of histidine. *ACS Chem. Neurosci.* 10:4729–4734.
39. Krimm, S., and Y. Abe. 1972. Intermolecular interaction effects in the amide I vibrations of polypeptides. *Proc. Natl. Acad. Sci. USA.* 69:2788–2792.
40. Ding, B., J. E. Laaser, ..., Z. Chen. 2013. Site-specific orientation of an  $\alpha$ -helical peptide ovispirin-1 from isotope-labeled SFG spectroscopy. *J. Phys. Chem. B.* 117:14625–14634.
41. Strasfeld, D. B., Y. L. Ling, ..., M. T. Zanni. 2009. Strategies for extracting structural information from 2D IR spectroscopy of amyloid: application to islet amyloid polypeptide. *J. Phys. Chem. B.* 113:15679–15691.
42. Hamm, P., and S. Woutersen. 2002. Coupling of the amide I modes of the glycine dipeptide. *Bull. Chem. Soc. Jpn.* 75:985–988.
43. Carr, J. K., A. V. Zabuga, ..., J. L. Skinner. 2014. Assessment of amide I spectroscopic maps for a gas-phase peptide using IR-UV double-resonance spectroscopy and density functional theory calculations. *J. Chem. Phys.* 140:224111.
44. Torii, H., and M. Tasumi. 1992. Model calculations on the amide-I infrared bands of globular proteins. *J. Chem. Phys.* 96:3379–3387.
45. Torii, H., and M. Tasumi. 1998. Ab initio molecular orbital study of the amide I vibrational interactions between the peptide groups in di- and tripeptides and considerations on the conformation of the extended helix. *J. Raman Spectrosc.* 29:81–86.
46. Lomont, J. P., J. S. Ostrander, ..., M. T. Zanni. 2017. Not all  $\beta$ -sheets are the same: amyloid infrared spectra, transition dipole strengths, and couplings investigated by 2D IR spectroscopy. *J. Phys. Chem. B.* 121:8935–8945.
47. Dunkelberger, E. B., M. Grechko, and M. T. Zanni. 2015. Transition dipoles from 1D and 2D infrared spectroscopy help reveal the secondary structures of proteins: application to amyloids. *J. Phys. Chem. B.* 119:14065–14075.

48. Gorkovskiy, A., K. R. Thurber, ..., R. B. Wickner. 2014. Locating folds of the in-register parallel  $\beta$ -sheet of the Sup35p prion domain infectious amyloid. *Proc. Natl. Acad. Sci. USA*. 111:E4615–E4622.
49. Choi, J.-H., S. Ham, and M. Cho. 2003. Local amide I mode frequencies and coupling constants in polypeptides. *J. Phys. Chem. B*. 107:9132–9138.
50. Ham, S., S. Cha, ..., M. Cho. 2003. Amide I modes of tripeptides: hessian matrix reconstruction and isotope effects. *J. Chem. Phys.* 119:1451–1461.
51. Ham, S., J.-H. Kim, ..., M. Cho. 2003. Correlation between electronic and molecular structure distortions and vibrational properties. II. Amide I modes of NMA–nD<sub>2</sub>O complexes. *J. Chem. Phys.* 118:3491–3498.
52. Buchete, N.-V., R. Tycko, and G. Hummer. 2005. Molecular dynamics simulations of Alzheimer's beta-amyloid protofilaments. *J. Mol. Biol.* 353:804–821.
53. Paravastu, A. K., A. T. Petkova, and R. Tycko. 2006. Polymorphic fibril formation by residues 10-40 of the Alzheimer's  $\beta$ -amyloid peptide. *Biophys. J.* 90:4618–4629.
54. Kim, Y. S., L. Liu, ..., R. M. Hochstrasser. 2009. 2D IR provides evidence for mobile water molecules in  $\beta$ -amyloid fibrils. *Proc. Natl. Acad. Sci. USA*. 106:17751–17756.
55. Møller, K. B., R. Rey, and J. T. Hynes. 2004. Hydrogen bond dynamics in water and ultrafast infrared spectroscopy: a theoretical study. *J. Phys. Chem. A*. 108:1275–1289.
56. Kwak, K., S. Park, ..., M. D. Fayer. 2007. Frequency-frequency correlation functions and apodization in two-dimensional infrared vibrational echo spectroscopy: a new approach. *J. Chem. Phys.* 127:124503.
57. Salehi, S. M., D. Koner, and M. Meuwly. 2019. Vibrational spectroscopy of N<sub>3</sub><sup>-</sup> in the gas and condensed phase. *J. Phys. Chem. B*. 123:3282–3290.
58. Desmond, J. L., D. Koner, and M. Meuwly. 2019. Probing the differential dynamics of the monomeric and dimeric insulin from amide-I IR spectroscopy. *J. Phys. Chem. B*. 123:6588–6598.
59. Torres, J., A. Kukol, ..., I. T. Arkin. 2001. Site-specific examination of secondary structure and orientation determination in membrane proteins: the peptidic (13)C=(18)O group as a novel infrared probe. *Biopolymers*. 59:396–401.

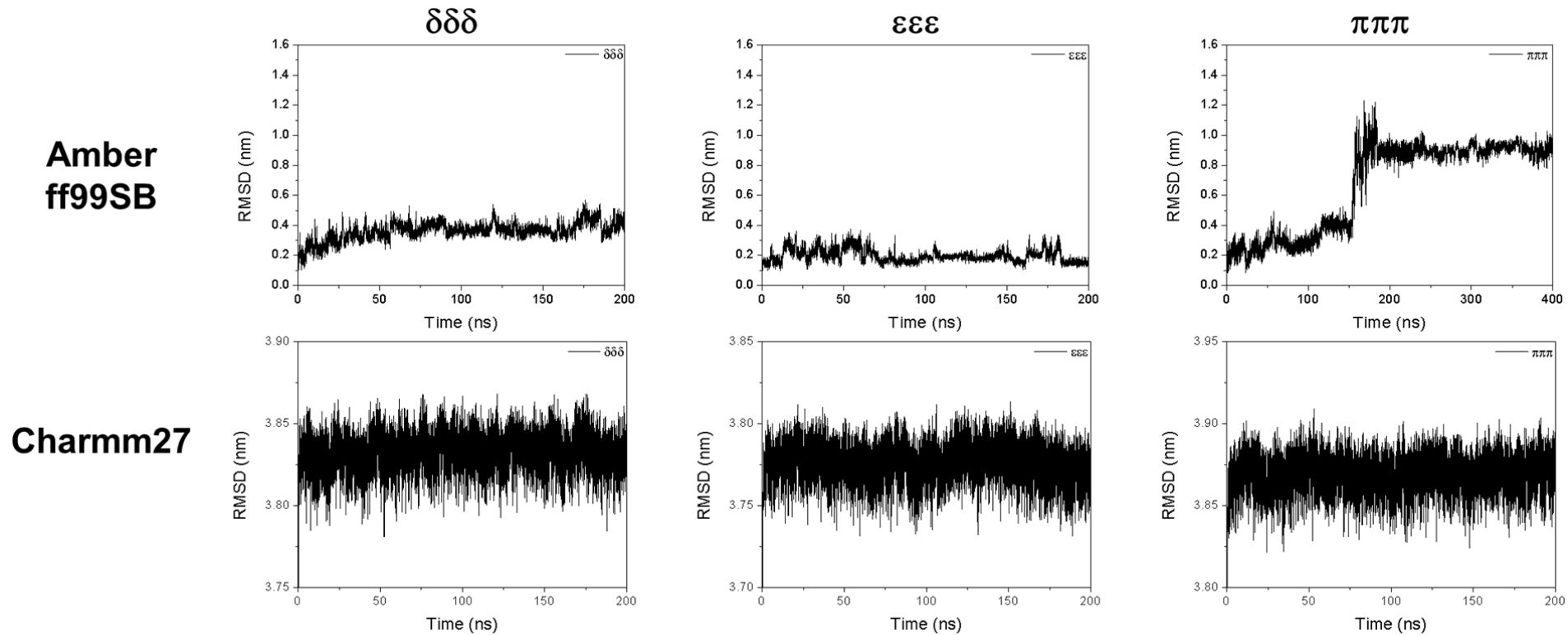


**Biophysical Journal, Volume 119**

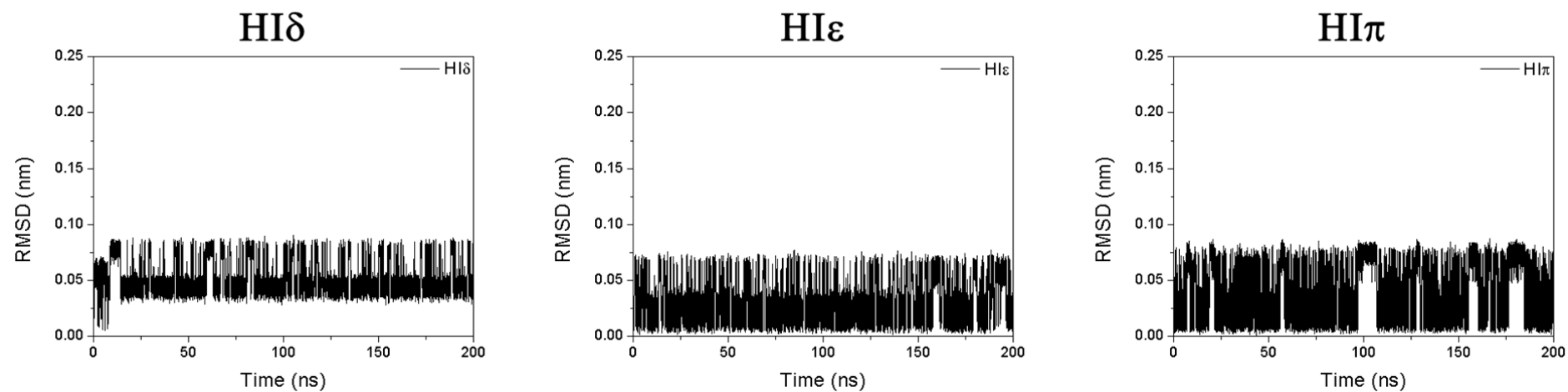
**Supplemental Information**

**Tautomeric Effect of Histidine on  $\beta$ -Sheet Formation of Amyloid Beta 1–40: 2D-IR Simulations**

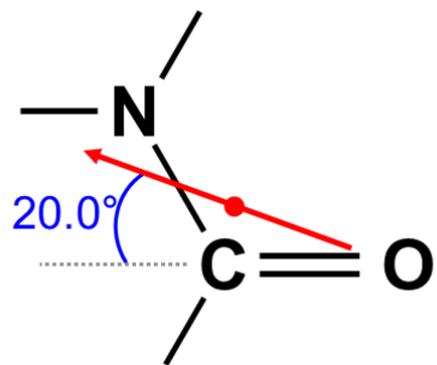
**Yeonsig Nam, Mahroof Kalathingal, Shinji Saito, and Jin Yong Lee**



**Figure S1.** Root mean square deviation (RMSD) of the  $\delta\delta\delta$ ,  $\epsilon\epsilon\epsilon$ , and  $\pi\pi\pi$  tautomers calculated with Amber ff99SB and Charmm27 force field.



**Figure S2.** Root mean square deviation (RMSD) of the HI $\delta$ , HI $\epsilon$ , and HI $\pi$  dipeptides calculated with Amber ff99SB force field.



**Figure S3.** Schematic drawing of the location (red circle) and direction (red arrow) of the transition dipole used in this study.







πππ

0.33	0.60	0.73	1.11	1.36	1.66	1.98	2.22	2.14	2.01	2.15	2.38	2.09	2.02	1.88	1.58	1.41	1.29	1.15	1.41	1.59	1.75	1.92	2.05	2.17	2.37	2.48	2.39	2.06	1.78	1.84	1.61	1.68	1.50	1.49	1.41	1.07	0.95
0.00	0.36	0.49	0.85	1.13	1.44	1.77	1.97	1.86	1.72	1.88	2.09	1.79	1.74	1.63	1.35	1.24	1.19	1.11	1.42	1.62	1.74	1.89	2.06	2.19	2.35	2.49	2.44	2.12	1.85	1.84	1.56	1.56	1.32	1.27	1.13	0.77	0.64
0.00	0.00	0.40	0.71	1.06	1.39	1.74	1.90	1.75	1.55	1.67	1.87	1.55	1.47	1.31	1.03	0.91	0.93	0.93	1.29	1.54	1.65	1.75	1.95	2.12	2.25	2.36	2.32	2.00	1.74	1.66	1.34	1.32	1.06	1.04	0.87	0.52	0.49
0.00	0.00	0.00	0.39	0.69	1.01	1.36	1.56	1.47	1.30	1.44	1.68	1.41	1.34	1.27	0.98	0.95	0.89	0.87	1.16	1.36	1.41	1.55	1.77	1.90	2.01	2.18	2.18	1.89	1.67	1.58	1.25	1.17	0.89	0.80	0.74	0.48	0.59
0.00	0.00	0.00	0.00	0.38	0.69	1.05	1.20	1.11	0.93	1.06	1.31	1.09	1.02	1.04	0.80	0.91	0.92	0.98	1.23	1.41	1.38	1.49	1.76	1.88	1.93	2.13	2.20	1.94	1.77	1.61	1.25	1.05	0.72	0.51	0.48	0.44	0.69
0.00	0.00	0.00	0.00	0.00	0.33	0.69	0.89	0.90	0.81	0.97	1.25	1.14	1.10	1.24	1.04	1.20	1.16	1.20	1.36	1.46	1.35	1.49	1.77	1.83	1.86	2.11	2.23	2.01	1.89	1.73	1.41	1.16	0.88	0.62	0.73	0.79	1.00
0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.64	0.78	0.78	0.90	1.19	1.19	1.16	1.37	1.23	1.42	1.37	1.41	1.49	1.54	1.37	1.52	1.80	1.82	1.80	2.10	2.25	2.08	2.01	1.84	1.54	1.26	1.03	0.77	0.95	1.09	1.31
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.72	0.87	1.00	1.25	1.35	1.37	1.64	1.55	1.77	1.73	1.77	1.81	1.81	1.61	1.76	2.03	2.01	1.96	2.29	2.48	2.34	2.30	2.13	1.86	1.57	1.37	1.10	1.27	1.42	1.62
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.66	0.80	0.96	1.14	1.24	1.59	1.57	1.86	1.90	1.99	2.08	2.12	1.92	2.01	2.31	2.32	2.23	2.54	2.75	2.61	2.58	2.36	2.06	1.73	1.49	1.19	1.28	1.49	1.69
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39	0.64	0.72	0.84	1.02	1.39	1.40	1.72	1.84	1.98	2.14	2.24	2.07	2.14	2.45	2.50	2.43	2.71	2.89	2.73	2.65	2.42	2.07	1.73	1.43	1.11	1.09	1.29	1.46	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31	0.47	0.50	0.64	1.00	1.04	1.38	1.54	1.72	1.90	2.04	1.90	1.92	2.24	2.34	2.26	2.50	2.67	2.50	2.41	2.15	1.78	1.44	1.13	0.82	0.78	1.06	1.30	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.47	0.48	0.87	0.96	1.32	1.48	1.69	1.85	1.99	1.83	1.80	2.12	2.24	2.13	2.34	2.54	2.38	2.32	2.02	1.67	1.31	1.04	0.78	0.83	1.19	1.47
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41	0.57	0.98	1.15	1.52	1.74	1.97	2.15	2.31	2.15	2.11	2.43	2.55	2.43	2.64	2.83	2.68	2.62	2.30	1.95	1.59	1.31	1.07	1.02	1.37	1.62
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.70	0.85	1.23	1.49	1.74	1.98	2.18	2.08	2.04	2.35	2.51	2.43	2.61	2.76	2.57	2.46	2.16	1.77	1.45	1.12	0.89	0.71	1.04	1.28	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41	0.59	0.97	1.23	1.49	1.71	1.92	1.82	1.75	2.06	2.24	2.15	2.30	2.45	2.27	2.17	1.84	1.47	1.14	0.84	0.68	0.61	1.01	1.31		
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.64	0.97	1.27	1.52	1.78	1.74	1.63	1.92	2.14	2.09	2.17	2.28	2.07	1.94	1.61	1.23	0.96	0.69	0.68	0.60	0.94	1.25	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.68	0.97	1.26	1.53	1.53	1.51	1.43	1.72	1.94	1.92	2.02	2.09	1.85	1.70	1.40	1.01	0.77	0.45	0.48	0.43	0.72	1.05	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.70	1.03	1.34	1.40	1.31	1.55	1.80	1.84	1.88	1.90	1.63	1.43	1.16	0.78	0.68	0.48	0.68	0.69	0.81	1.09	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.63	0.95	1.04	0.98	1.20	1.45	1.52	1.56	1.55	1.26	1.05	0.83	0.45	0.48	0.46	0.74	0.89	0.96	1.23				
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41	0.73	0.89	0.90	1.06	1.29	1.42	1.47	1.42	1.10	0.86	0.75	0.47	0.63	0.68	0.92	1.10	1.09	1.31		
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.54	0.57	0.67	0.89	1.05	1.10	1.04	0.73	0.54	0.49	0.42	0.66	0.87	1.12	1.39	1.43	1.67		
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31	0.47	0.48	0.60	0.82	0.91	0.87	0.61	0.54	0.61	0.69	0.87	1.11	1.31	1.63	1.68	1.91		
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.47	0.50	0.63	0.84	0.91	0.75	0.79	0.74	0.78	0.82	1.07	1.23	1.59	1.70	1.96		
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.52	0.54	0.65	0.76	0.63	0.71	0.53	0.61	0.68	1.01	1.23	1.60	1.76	2.05			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.46	0.44	0.46	0.46	0.37	0.57	0.54	0.79	0.96	1.30	1.53	1.89	2.02	2.29			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.47	0.55	0.58	0.81	0.86	1.08	1.19	1.50	1.69	2.07	2.19	2.44					
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39	0.66	0.78	1.03	0.96	1.14	1.18	1.50	1.67	2.07	2.25	2.53		
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.59	0.90	0.84	1.13	1.25	1.62	1.85	2.23	2.40	2.68			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.68	0.76	1.12	1.34	1.71	1.96	2.31	2.43	2.69			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.49	0.86	1.13	1.48	1.74	2.06	2.15	2.39			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.71	1.05	1.35	1.62	1.90	1.94	2.15			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.72	1.06	1.36	1.66	1.77	2.03			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.67	0.99	1.26	1.38	1.66		
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39	0.68	1.01	1.22	1.54			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.62	0.86	1.20			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41	0.71	1.05			
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41	0.74			
0.00	0.00	0.00</																																			

**Table S2.** The Amide-I Vibrational Frequency of the HI $\delta$ , HI $\epsilon$  and HI $\pi$  for Histidine Residue, Histidine Residue with Solvent, and Dipeptide System Calculated with MP2/6-31G\*\*.

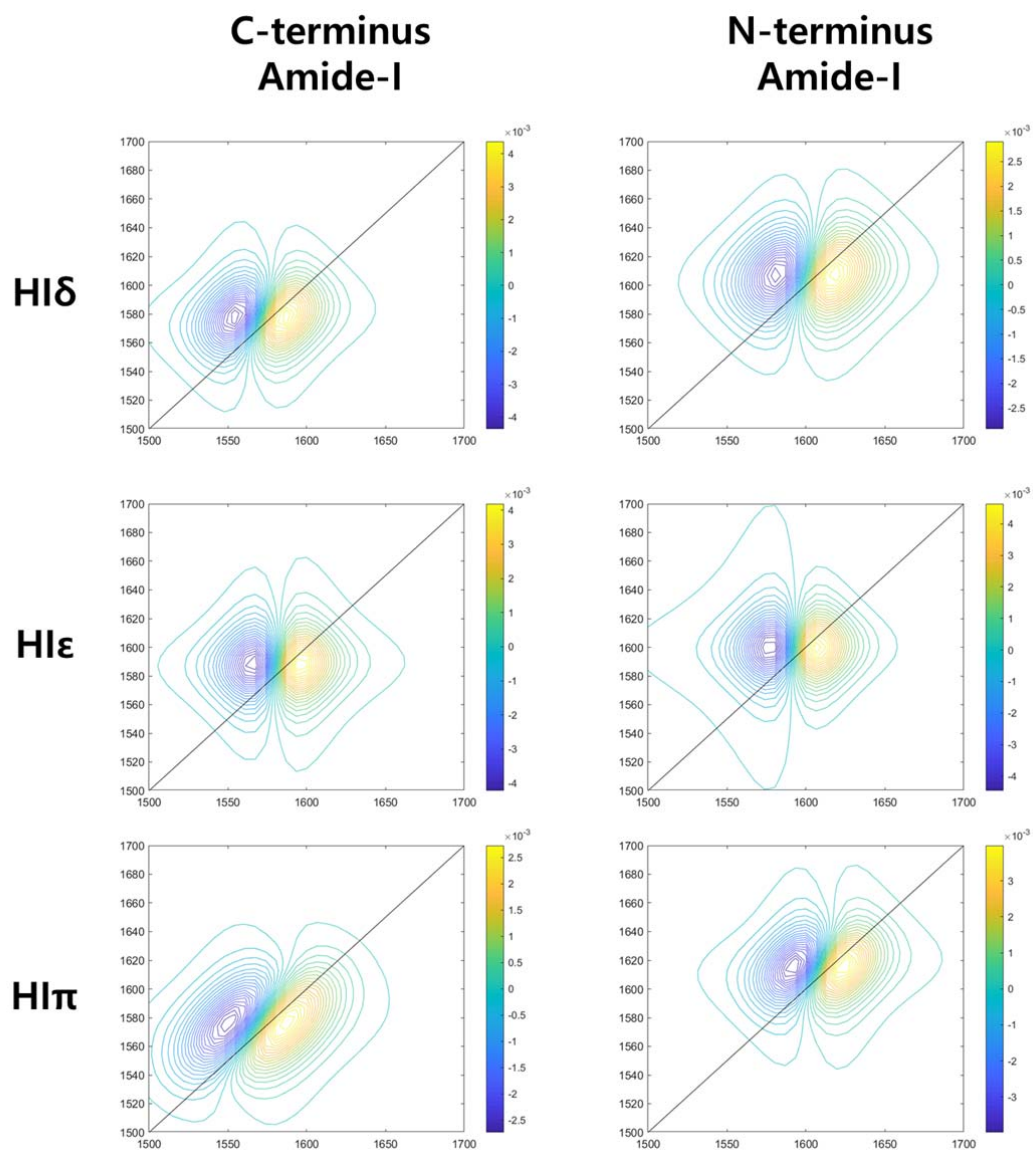
	HI $\delta$	HI $\epsilon$	HI $\pi$
Histidine residue	1736.303	1726.933	1738.286
Histidine residue with implicit solvent	1721.418	1712.814	1712.108
Dipeptide (labeled)	1687.192	1691.288	1667.768
Dipeptide (non-labeled)	1767.503	1762.266	1758.551

**Table S3.** The Number of Hydrogen Bonds within 0.35 nm from amide-I bond for C/N-terminus Amide-I Bond for the HI $\delta$ , HI $\epsilon$  and HI $\pi$  Tautomers.

C-terminus amide-I bond		
HI $\delta$	HI $\epsilon$	HI $\pi$
2.967	2.054	2.533
N-terminus amide-I bond		
HI $\delta$	HI $\epsilon$	HI $\pi$
3.413	3.450	3.109

**Table S4.** The atomic charge of C and O atom for C-terminus and N-terminus amide-I bonds based on AMBERff99SB force field.

Atom type		C-terminus	N-terminus
		Amide-I bond	Amide-I bond
HI $\delta$	C	0.5973	0.5972
	O	-0.5679	-0.5679
HI $\epsilon$	C	0.5973	0.5972
	O	-0.5679	-0.5679
HI $\pi$	C	0.7341	0.5972
	O	-0.5894	-0.5679



**Figure S4.** 2DIR spectra for amide-I vibration of HI $\delta$ , HI $\epsilon$ , and HI $\pi$  dipeptides: C-terminus (left) and N-terminus (right) amide bond at waiting time,  $t_2 = 3$  ps.