

## Research article

Helix-to-sheet transition of the A $\beta$ 42 peptide revealed using an enhanced sampling strategy and Markov state modelHuilin Wen<sup>a,b</sup>, Hao Ouyang<sup>a</sup>, Hao Shang<sup>a</sup>, Chaohong Da<sup>a</sup>, Tao Zhang<sup>a,\*</sup><sup>a</sup> School of Biomedical Engineering and Technology, Tianjin Medical University, Tianjin 300070, PR China<sup>b</sup> The Third Hospital of Hebei Medical University, Shijiazhuang 050051, PR China

## ARTICLE INFO

## Keywords:

Enhanced sampling strategy  
ANN  
Metadynamics  
Markov state model  
Conformational transition  
A $\beta$  peptide

## ABSTRACT

The self-assembly of A $\beta$  peptides into toxic oligomers and fibrils is the primary cause of Alzheimer's disease. Moreover, the conformational transition from helix to sheet is considered a crucial step in the aggregation of A $\beta$  peptides. However, the structural details of this process still remain unclear due to the heterogeneity and transient nature of the A $\beta$  peptides. In this study, we developed an enhanced sampling strategy that combines artificial neural networks (ANN) with metadynamics to explore the conformational space of the A $\beta$ 42 peptides. The strategy consists of two parts: applying ANN to optimize CVs and conducting metadynamics based on the resulting CVs to sample conformations. The results showed that this strategy achieved better sampling performance in terms of the distribution of sampled conformations. The sampling efficiency is increased by 10-fold compared to our previous Hamiltonian Exchange Molecular Dynamics (MD) and by 1000-fold compared to ordinary MD. Based on the sampled conformations, we constructed a Markov state model to understand the detailed transition process. The intermediate states in this process are identified, and the connecting paths are analyzed. The conformational transitions in D23-K28 and M35-V40 are proven to be crucial for aggregation. These results are helpful in clarifying the mechanism and process of A $\beta$ 42 peptide aggregation. D23-K28 and M35-V40 can be identified as potential targets for screening and designing inhibitors of A $\beta$  peptide aggregation.

## 1. Introduction

As the most common age-associated disorder, Alzheimer's disease (AD) has become a serious social and health problem worldwide [1]. AD is pathologically characterized by the presence of Amyloid- $\beta$  (A $\beta$ ) plaques and Tau neurofibrillary tangles in the brain [2,3]. A $\beta$  plaques are composed of insoluble A $\beta$  fibrils and other aggregated forms of A $\beta$  peptides. The A $\beta$  peptide is produced from the cleavage of the amyloid precursor protein (APP) by secretases and has a variable length ranging from 36 to 43 residues [4]. Because the cleaved segment is primarily located in the transmembrane region of APP, the A $\beta$  peptide maintains a helical conformation after it is produced [5]. In addition, it has been confirmed that the A $\beta$  peptides exist in highly disordered states and only adopt a limited number of stable and ordered conformations under physiological conditions [6]. These ordered conformations can transform into each other through the disordered conformational states [7]. There is also a central hydrophobic core consisting of multiple hydrophobic residues in the middle portion of the A $\beta$  peptide. Consequently, the A $\beta$  peptides tend to form an intramolecular  $\beta$ -sheet structure that is

stabilized by hydrogen bonds and hydrophobic interactions [8]. A recent study has confirmed that a specific  $\beta$ -hairpin plays a crucial role in the formation of A $\beta$  oligomers [9]. Following the formation of intramolecular  $\beta$ -sheet structures, A $\beta$  peptides can further aggregate into oligomers and fibrils, which are highly cytotoxic to neurons [10]. The experimental and computational studies have confirmed that the presence of intramolecular  $\beta$ -sheet structure is important for the aggregation of A $\beta$  peptides and the formation of A $\beta$  fibers [11–14]. Furthermore, the studies have observed the formation of oligomeric intermediates containing the  $\alpha$ -helix in the C-terminus [15,16]. These  $\alpha$ -helical oligomers can further initiate amyloid aggregation by converting into partial amyloid-like  $\beta$ -structures [16]. As the self-assembly of the A $\beta$  peptides into toxic oligomers and fibrils is a key causal event in the onset of AD [17], it is crucial to gain a deeper understanding of the conformational changes from helix to sheet that occur in both A $\beta$  monomers and oligomers. This understanding is essential for investigating amyloid aggregation and designing inhibitors.

The highly aggregated structures of A $\beta$  peptides have been characterized and determined directly using experimental methods [18].

\* Corresponding author.

E-mail address: [zhangtao@tmu.edu.cn](mailto:zhangtao@tmu.edu.cn) (T. Zhang).

According to the identified structures, all types of A $\beta$  fibrils share a common core structure consisting of an array of intermolecular  $\beta$ -sheets that are stabilized by hydrogen bonds [19]. But the mechanism and process of A $\beta$  peptide aggregation are not clearly understood due to the heterogeneity and transient nature of A $\beta$  oligomeric species [20]. During the process of aggregation, A $\beta$  peptides that form sheet conformations have a strong tendency to aggregate. At the initial stage of aggregation, A $\beta$  peptides with a relatively high intramolecular  $\beta$ -sheet content tend to aggregate into dimers due to stable interactions between  $\beta$ -strands [21]. In the middle stage of aggregation, when adopting the intramolecular  $\beta$ -sheet conformation, A $\beta$  peptides are prone to binding with A $\beta$  prefibrils [22]. But the related experiments show that the A $\beta$  peptide is generated from the transmembrane segment of APP, and the helical portion of the transmembrane is located at the N-terminal of the A $\beta$  peptide [23,24]. Therefore, the A $\beta$  peptide must undergo a conformational transition from  $\alpha$ -helix to intramolecular  $\beta$ -sheet before aggregation. As the essential step for A $\beta$  aggregation, this conformational transition has also been the focus of extensive studies in recent years [25–28]. The investigation of conformational transitions can be helpful in probing the initial stages of aggregation and providing valuable structural insights into the inhibition of aggregation.

Compared to the physicochemical techniques used to study A $\beta$  peptides, molecular dynamics (MD) simulation can provide atomic-level spatial and temporal resolution for depicting the structure and dynamics of A $\beta$  peptides [29]. By employing the MD method, we can accurately describe the details of conformational transition of the A $\beta$  peptides. According to MD simulation results, A $\beta$  peptides exhibit a rugged free energy surface and the coexistence of multiple minima corresponding to different native conformations [30]. Moreover, A $\beta$  peptides not only adopt stable, well-defined secondary or tertiary structures at equilibrium, but also dynamically interconvert between various conformations [31]. These properties may further extend the application of MD to study the conformations of A $\beta$  peptides [19]. In the meantime, these characteristics also posed more difficulties in capturing key conformations, thereby limiting the application of MD in investigating conformational transitions of the A $\beta$  peptides [29]. In order to effectively explore the conformation space of the A $\beta$  peptides, an enhanced sampling strategy is required to improve sampling efficiency. When obtaining a sufficient number of conformations of the A $\beta$  peptides, it becomes possible to gain a comprehensive understanding of the conformational transition process in relation to the kinetic properties of A $\beta$  peptide conformations [32].

In this study, we specifically investigated the transition process in which the A $\beta$  peptide changes from  $\alpha$ -helix to intramolecular  $\beta$ -sheet state. Metadynamics, an efficient enhanced simulation, was used to perform biased simulations. This was achieved by introducing external Gaussian potentials that act on specific collective variables (CVs) that accurately describe the transition process. An artificial neural network was constructed and trained to extract the correct CVs from the MD trajectory, which undergoes similar conformational transitions. In total, 10 iterations of metadynamics were run, producing a 300 ns MD trajectory. By comparing our MD trajectory of the A $\beta$  peptide with other simulations, we found that the conformation distribution extracted from our nanosecond-scale simulation is comparable to those obtained from microsecond-scale simulations. This indicates that our sampling efficiency has significantly improved using this strategy. Subsequently, we constructed a Markov state model (MSM) using the sampled conformations. Based on the model, we depicted the conformational transition and identified three crucial regions that play an important role in this process. We believe that the results can provide a more comprehensive understanding of the atomic-level transformation of the A $\beta$  peptide from helix to intramolecular sheet state.

## 2. Methods

Among all A $\beta$  isoforms, A $\beta$ 40 and A $\beta$ 42 are the most important ones [33]. Moreover, A $\beta$ 42 is more prone to aggregation than A $\beta$ 40 due to the

presence of two additional hydrophobic residues [34]. Therefore, the A $\beta$ 42 structures with helical conformations were used as the initial conformations in this study. Two NMR structures (PDB codes 1IYT and 1Z0Q) were directly obtained from the Protein Data Bank and considered to represent the helix conformation of A $\beta$ 42 in a solution environment. The third structure was generated by combining two structures (PDB codes 2BP4 and 2LP1). These structures represent the solution structures of the N-terminal 1–16 segment of A $\beta$ 42 and the transmembrane C-terminal domain of the amyloid precursor protein, respectively. The N-terminal region of 2BP4 and the C-terminal region of 2LP1 are conserved and joined together to form a single structure. This spliced structure can be considered as the helix conformation released from APP, and consequently it was named APP. Three A $\beta$ 42 structures are shown in Fig. 1, along with their helical content.

Metadynamics is an efficient accelerated sampling method for observing rare events that cannot be accessed using conventional simulation methods [35]. In metadynamics, rare events can be sampled by applying a biased potential to the collective variables (CVs) under consideration and pushing the simulated system away from local free energy minima. Therefore, it is necessary and crucial to define the appropriate CVs that can accurately describe the dynamics of the simulated system when performing metadynamics. In this study, we employ a neural network to extract and optimize the appropriate CVs from the obtained MD trajectory. The resulting CVs are then used to guide metadynamics simulations.

### 2.1. Determination and optimization of CVs

Anncolvar is a specific tool for approximately extracting CVs using machine learning from MD trajectory data [36]. The output is the optimized combination of CVs that can describe the time-dependent structural evolution. These optimized CVs are organized into a file to be used as input for running metadynamics simulations. In this study, we developed a strategy that combines Anncolvar with metadynamics. We trained a multi-layer network using the simulation trajectory from previous experiments to generate the appropriate CVs. The metadynamics simulations were then conducted using these CVs. Moreover, the resulting metadynamics simulation trajectories were reintroduced into the constructed network for further optimization of CVs. Fig. 2 shows the flowchart of the derivation of CVs and the basic architecture of the network.

As depicted in the flowchart, the A $\beta$ 42 structures in the chosen MD trajectories were positioned at the center of a 20 nm box. Additionally, the translation of individual atoms in the structure was eliminated to ensure a more precise alignment of each structure. The processed 3D coordinates were continuously input into the networks, which comprised of three hidden layers. There are 256, 128, and 64 neurons in each layer, respectively. The tanh activation function is used. After performing efficient dimensional reduction, several characteristics were extracted and combined as the CVs for running metadynamics simulations.

In the previous research, we simulated the conformational transition of the A $\beta$ 42 peptide from helix to sheet using Hamiltonian simulation [37]. The trajectories undergoing significant conformational fluctuations were selected to train the multi-layer network and extract the initial CVs. We first compared the root mean square deviation (RMSD) of the A $\beta$ 42 peptide in four 400 ns MD trajectories for 1IYT, 1Z0Q, and APP, respectively. Based on the comparison shown in Supplement Fig. 1, the trajectory with the highest RMSD fluctuation was used to train the neural network model. Additionally, we calculated the root mean square fluctuation (RMSF) for each residue of the A $\beta$ 42 peptide in the chosen trajectory. Fig. 3 shows three disconnected regions, namely K16-V18, A21-D23, and K28-A30, displaying significant structural fluctuations. The results indicate that these regions are likely to be involved in significant structural fluctuations. Although the F19-F20 region has a similar RMSF to those three regions, it is not considered due to its

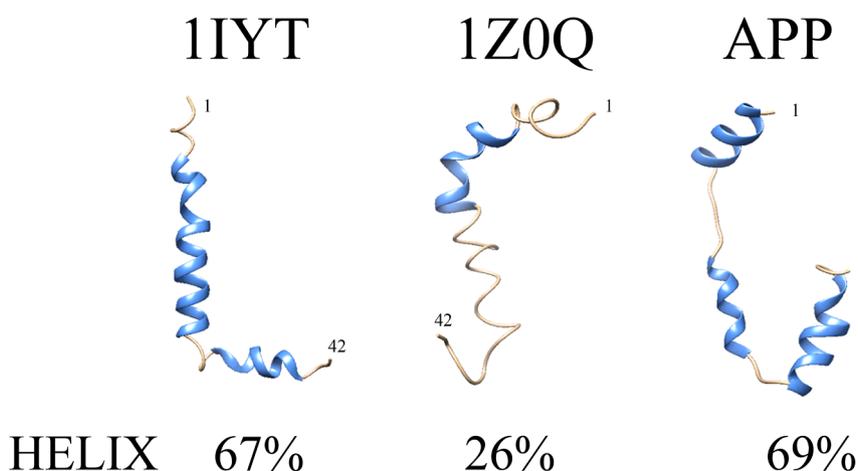


Fig. 1. Three initial Aβ42 structures and their corresponding helical content.

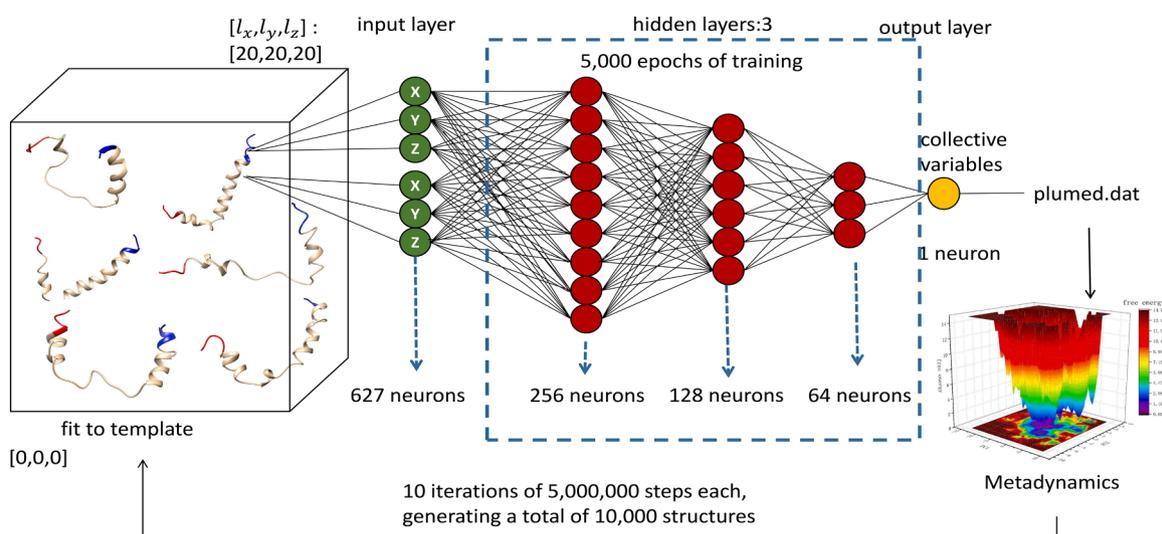


Fig. 2. The flowchart of CVs extraction and optimization using neural networks. The atomic coordinates of the Aβ42 peptide in various conformations were sequentially input into the neural networks. After performing efficient dimensional reduction, several characteristics were extracted and combined as the CVs for running metadynamics simulations.

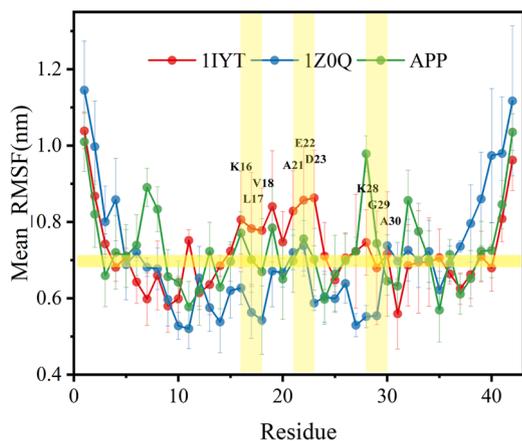


Fig. 3. Comparison of individual residue RMSF in Aβ42 peptide. A threshold RMSF of 0.7 nm is set, above which the structure of the corresponding region fluctuates more. According to this threshold, three disconnected regions close to the center of the Aβ42 peptide, namely K16-V18, A21-D23 and K28-A30, exhibit significant structural fluctuation.

connection to A21-D23 and its proximity to K16-V28.

Therefore, the distance between adjacent residues in these regions and their dihedral angles are considered important characteristics of conformational changes. With the selected trajectory and relevant information, we trained the neural network constructed by Anncolvar and obtained the original CVs which are a combination of atomic coordinates from the selected regions. The original CVs can potentially be used to guide the metadynamics. The newly generated MD trajectory was once again input into the neural network and subsequently obtained the optimized CVs with different weighting efficiency compared to the original CVs. The optimized CVs were used to direct metadynamics. A total of 10 rounds of metadynamics were conducted with the CVs continuously optimized using the snapshots from the previous cycle.

### 2.2. Metadynamics

Three initial structures of the Aβ42 peptide were solvated with TIP3P water molecules and placed in a cubic box with a boundary 1.0 nm away from the edges. The MD parameters were set up the same as in the previous study [37]. The parameters associated with the Gaussian potential are defined as follows: the initial height of the potential is set to 0.6 kJ/mol and is adjusted every 500 time steps. The analysis uses a

Gaussian width of 0.35, which is generally of the same order of magnitude as the standard deviation of the CVs in the unbiased run. The bias factor, as an important parameter, should be carefully chosen because it determines the effectiveness of the sampling. After several trials, the bias factor was finally selected as 6.0, and the value of SIGMA is 0.35.

In order to comprehensively sample the conformational space of the A $\beta$  peptide, multiple rounds of metadynamics simulations were conducted. These simulations utilized different CVs that were optimized based on the continuously generated simulation trajectory. Ten rounds of metadynamics simulations were conducted for three A $\beta$  peptides, resulting in a total simulation trajectory of 300 ns. All simulations were performed with the OPLS-AA/L force field using the PLUMED 2.6 package and Gromacs 2018.8 [38].

### 2.3. Construction of MSM

All A $\beta$  peptide conformations sampled by the metadynamics simulation were further analyzed to explore the transition process among different conformations using Markov state modeling (MSM). In the MSM model, the conformation transition can be modeled as a collection of discrete Markov jump processes [39]. In order to clearly illustrate this process, MSM is generally also viewed as a network. The network nodes represent different conformation states, and the edges indicate the probabilities of transitioning between two states. When estimating the probability of a specific transition using an MSM model, we do not require information about previously visited conformations.

In this study, we compared five features that represent the characteristics of secondary structures. We calculated their VAMP-2 scores at a specific lag time and selected the feature with the highest VAMP-2 score to reduce dimensionality using time-lagged independent component analysis (tICA). In the lower-dimensional space, all conformations of the A $\beta$  peptide were clustered into distinct microstates based on their structural similarity. The microstates were then combined into several macrostates using the Robust Perron Cluster Analysis (PCCA+) algorithm, which considers only their kinetic similarities. From each macrostate, we extracted representative conformations to illustrate the transition process. The transition matrix was also constructed based on the resulting network. According to the transition matrix, we have identified the primary transition pathway. All the above analyses were performed using Python 3.7 and PyEMMA package 2.5.5 [40].

## 3. Results and discussion

### 3.1. Conformation sampling by metadynamics

#### 3.1.1. Conformational comparison of A $\beta$ 42 peptide

In order to clearly display the sampling results of metadynamics, we projected the sampled conformations onto the subspace defined by the secondary structure contents. Supplement Fig. 2 shows the distribution of sampling conformations in the subspace comprising Coil and Sheet contents for 1IYT, 1Z0Q, and APP. Based on the conformational distribution, we observe that the sampling range continuously increases as the simulation progresses. This expansion of the sample space is attributed to the utilization of different CVs by multiple-round optimization. The optimized CVs with varying weights can enhance the crossing of energy barriers and sample different conformations during metadynamics.

Taking all sampled conformations into consideration, the predominant conformations of 1IYT are concentrated in the region characterized by an intramolecular sheet content of 10–30% and a coil content of 20–70%. But the sampled conformations of 1Z0Q and APP are concentrated in the range of 30% to 80% for coil content and 10% to 50% for sheet content. Compared to the conformations sampled from 1IYT, these structures contain more sheet or coil formations. In addition, the distribution curves of the sampled conformations partially overlap across

the different rounds of simulation, indicating that the entire sampling process is continuous and convergent.

Similarly, the same trend is also observed in the distribution within the subspace consisting of helix and sheet contents, as shown in Supplement Fig. 3. The major sampled conformations of 1IYT are distributed within the range of 20–70% for helix content and 10–30% for sheet content. The conformations sampled from 1Z0Q and APP primarily fall within the range of 10–60% for helix content and 10–50% for sheet content. The increase in sheet content, accompanied by a decrease in helix content, indicates a conformational transition from helix to sheet.

All sampled conformations for 1IYT, 1Z0Q, and APP were combined and visualized in the subspace defined by the secondary structure contents. As shown in Fig. 4, the sampled conformations of the A $\beta$  peptide occupied almost the entire range of subspace, except for the upper right corner where the content of secondary structure exceeded 60%. There are no naturally occurring conformations of the A $\beta$  peptide that adopt such a high content of secondary structures. Therefore, we believe that the sampled conformations encompass most of the possible natural conformations of the A $\beta$  peptide.

Additionally, a clustering analysis was performed on all sampled conformations to assess their structural diversity. If the RMSD of two sampled conformations is within 3 Å, these two conformations will be classified into the same cluster. The results are shown in Supplement Fig. 4. It can be observed that the number of structural clusters gradually increases as the sampling progresses. The increased number of clusters implies that the sampled conformations exhibit greater structural diversity. Moreover, compared to the initial several rounds of sampling, the rate of increase gradually decreases in the later rounds of sampling. It indicates that the number of newly sampled conformations is decreasing, and the sampling tends to converge to the maximum value.

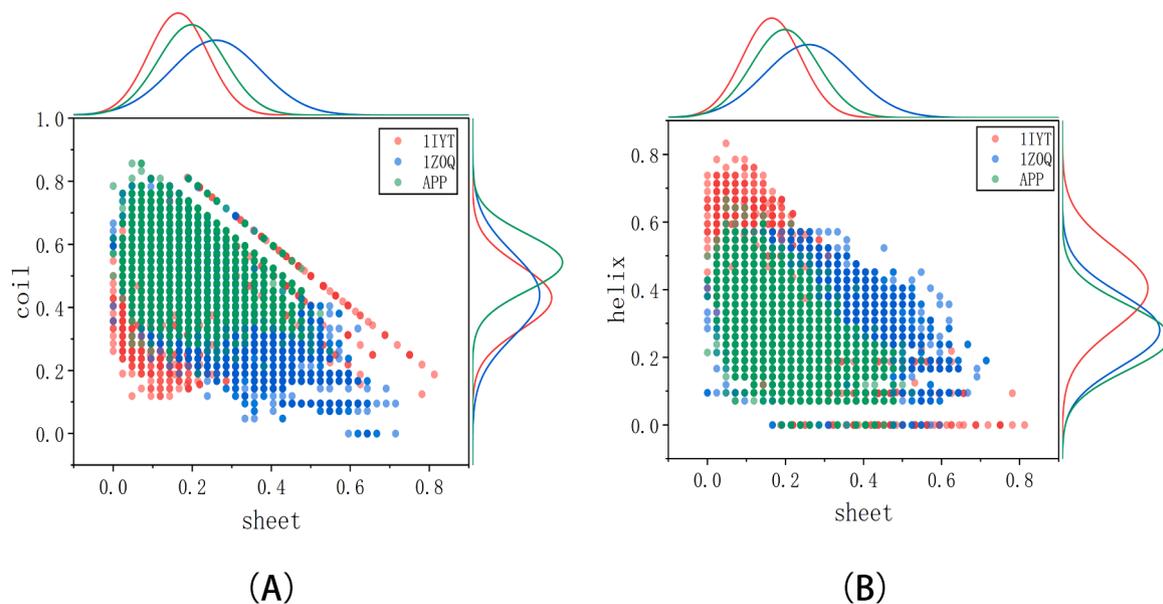
We further calculated the secondary structure contents of A $\beta$  peptides in the fibrillar, oligomeric, and monomeric states stored in the PDB database. The structural contents of each A $\beta$  peptide are listed in Supplement Table 1. These A $\beta$  peptide structures were projected onto the same subspace as the sampled conformations. As shown in Fig. 5, the experimentally determined structures of the A $\beta$  peptide are also within the sampled range. It indicates that the sampled conformations include various states of A $\beta$  peptides that occur in nature, and these conformations are suitable for analyzing conformational transitions.

The conformational distribution displays a comprehensive sampling of the A $\beta$  peptide, including the conformations of monomers, oligomers, and fibrils. Moreover, the results of the clustering analysis show that the newly sampled conformations are similar to the original conformations. These results demonstrate that convergence was essentially achieved after 10 rounds of metadynamics simulations.

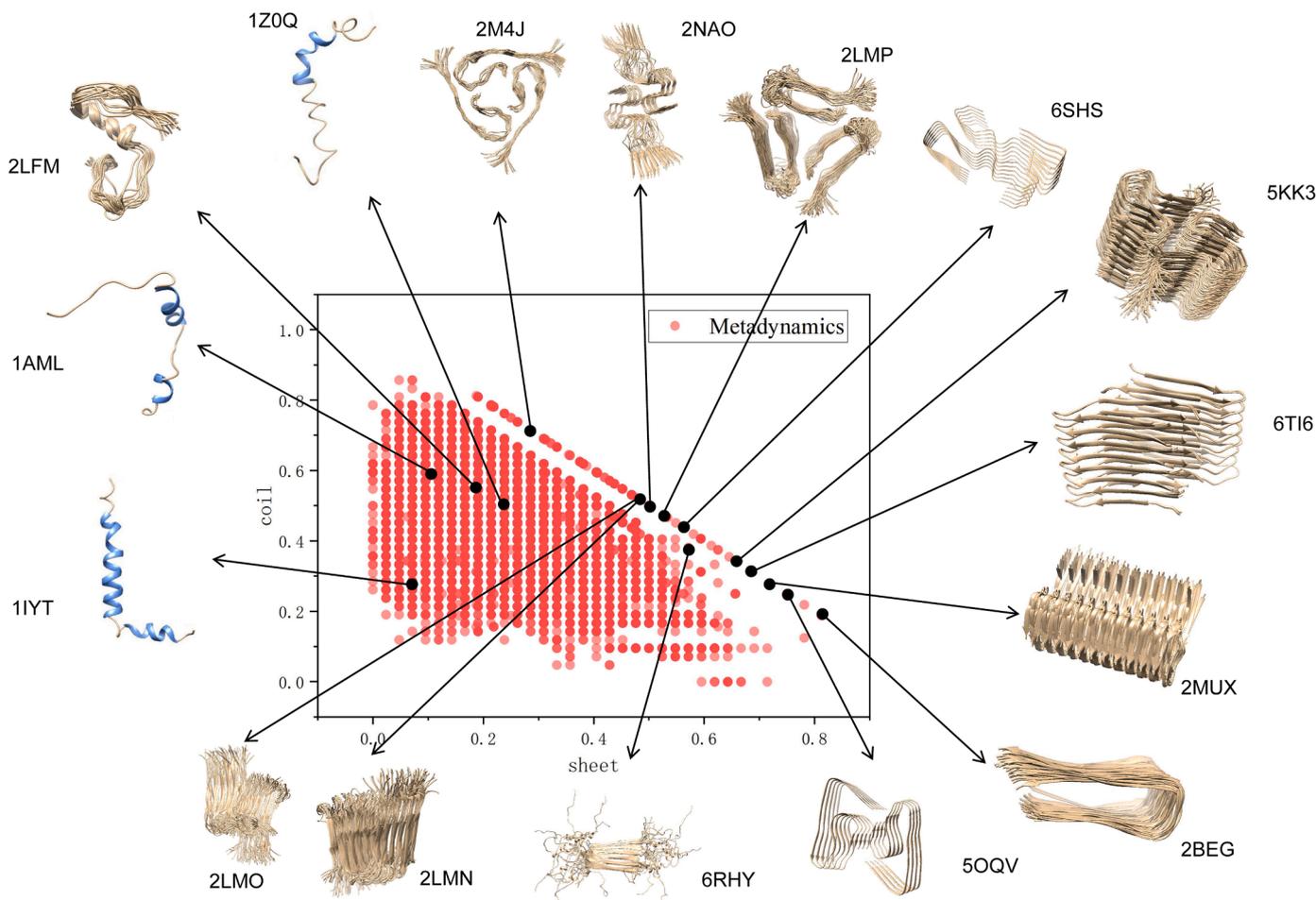
#### 3.1.2. Comparison of search efficiency among different simulations

Finally, we measured the sampling efficiency of the multiple rounds of metadynamics by comparing the conformational distribution with two other simulations related to the A $\beta$  peptide: the previous 4.8  $\mu$ s Hamilton replica-exchange molecular dynamics (H-REMD) [37] and the 315  $\mu$ s conventional molecular dynamics simulation [7]. Fig. 6 displays the conformational distributions in the subspace defined by the helix and sheet content for three simulations.

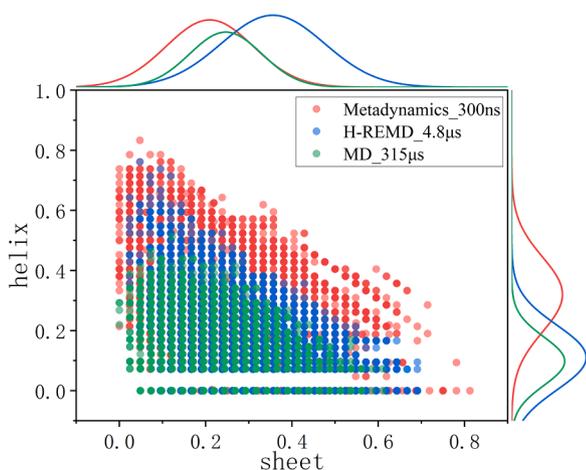
We observe that the entire range of sampling partially overlaps in three simulations. According to the comparison of distribution curves, we find that the range of the curve for multiple rounds metadynamics essentially encompasses the other two curves. But the peak of each curve is slightly different from one another. The conformations generated by multiple rounds of metadynamics mainly concentrate on a region with a helix content of 30% and an intramolecular sheet content of 25%. Whereas the majority of conformations generated by the previous H-REMD simulation had a helical content of 10% and a sheet content of 35%. Most conformations sampled by the conventional molecular dynamics simulations have similar helix and sheet content as the H-REMD simulation and multiple rounds of metadynamics, respectively.



**Fig. 4.** The distribution of sampled conformations of 1IYT, 1Z0Q, and APP on the subspace defined by Coil and Sheet contents (A) and Helix and Sheet contents (B). Distribution curves of sampled conformations are shown at the top and right side of the figure. These distribution curves indicate that the sampling ranges partially overlapped for 1IYT, 1Z0Q, and APP.



**Fig. 5.** The distribution of Aβ structures in different states, obtained from the PDB database, is projected onto the subspace defined by coil and sheet content. In this subspace, the sampled conformations are distributed. Based on the distribution, we observe that all Aβ structures in different states fall within the sampling range in this study.



**Fig. 6.** Comparison of conformational distributions from three simulations. Distribution curves of sampled conformations are shown at the top and right sides of the figure. These distribution curves indicate that the sampling ranges partially overlapped for three simulations.

According to the overlapping distribution, we can estimate the sampling efficiency by calculating the ratio of two simulated times. Compared to conventional simulation, the sampling efficiency is significantly improved by about 1000-fold through the utilization of multiple rounds

of metadynamics. Compared to H-REMD, the sampling efficiency is increased by about 16-fold. The adopted strategy of combining ANN with metadynamics can significantly enhance the sampling efficiency. But the strategy also has limitations. For example, the sampling efficiency depends on the initial CVs, the configuration of ANN structure, and the optimization of network parameters. The comparison of sampling efficiency using different setups will be performed in future research to develop an appropriate procedure for protein analysis.

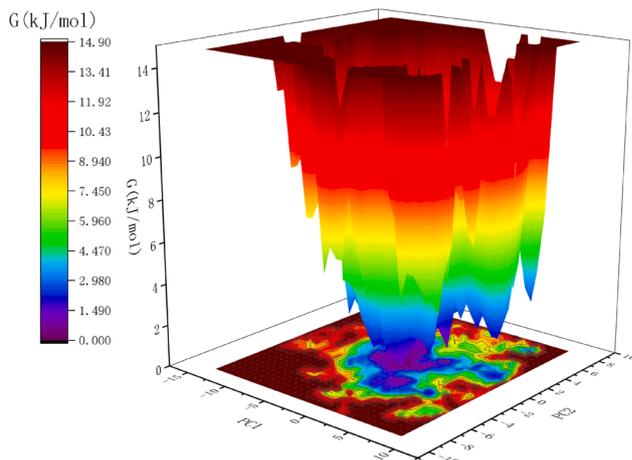
### 3.2. Conformational analysis of Aβ42 peptide sampled by metadynamics

#### 3.2.1. Principal component analysis of sampled conformations

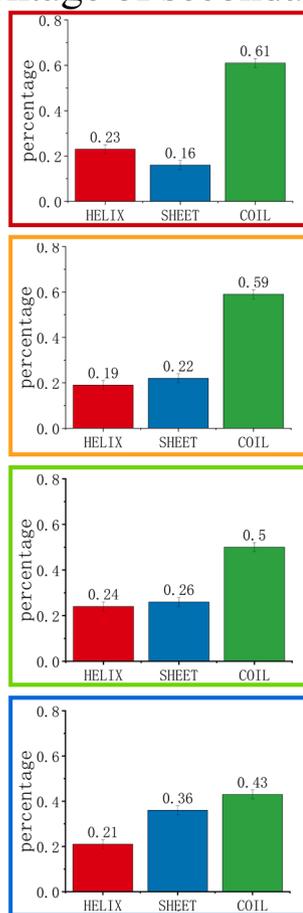
Based on the above comparison of Aβ peptide conformations sampled by metadynamics, it is evident that the sampled conformations are more diverse and covered almost the entire range defined by various secondary structures. In order to identify the significant conformational differences, we conducted a principal component analysis (PCA) on all sampled Aβ peptide conformations.

All sampled Aβ peptide conformations from 1IYT, 1Z0Q, and APP are being compared to one another. The eigenvectors and corresponding eigenvalues are calculated based on the variance of the structure. Supplement Fig. 5 shows the spectrum of eigenvalues that represent the structural variance captured by each principal component (PC). The first three components, PC1-PC3, can only account for half of the structural variance in all conformations. It is consistent with the clustering result that the sampled Aβ peptide conformations are more diverse.

## Free-energy landscape map



## Percentage of secondary structure



**Fig. 7.** Free energy map of the sampled conformations and comparison of secondary structure contents calculated from the conformations located in four regions with different free energy gradients. The color of frame borders around secondary structure contents is the same as that of one of the regions. Comparison of secondary structure contents indicates that Aβ peptide conformations with higher energy are typically in a disordered state, possessing a higher coil content. On the other hand, conformations located in the lower energy regions mainly adopt helix or sheet structures.

Subsequently, we attempt to project the conformations onto lower-dimensional subspaces defined by these three PCs. But we find that the sampled conformations cannot be clearly classified into distinct groups, as shown in Supplement Fig. 6. It may be due to the fact that the A $\beta$  peptide does not form a stable tertiary structure as a whole, and the major structural fluctuations are likely to occur within the local regions.

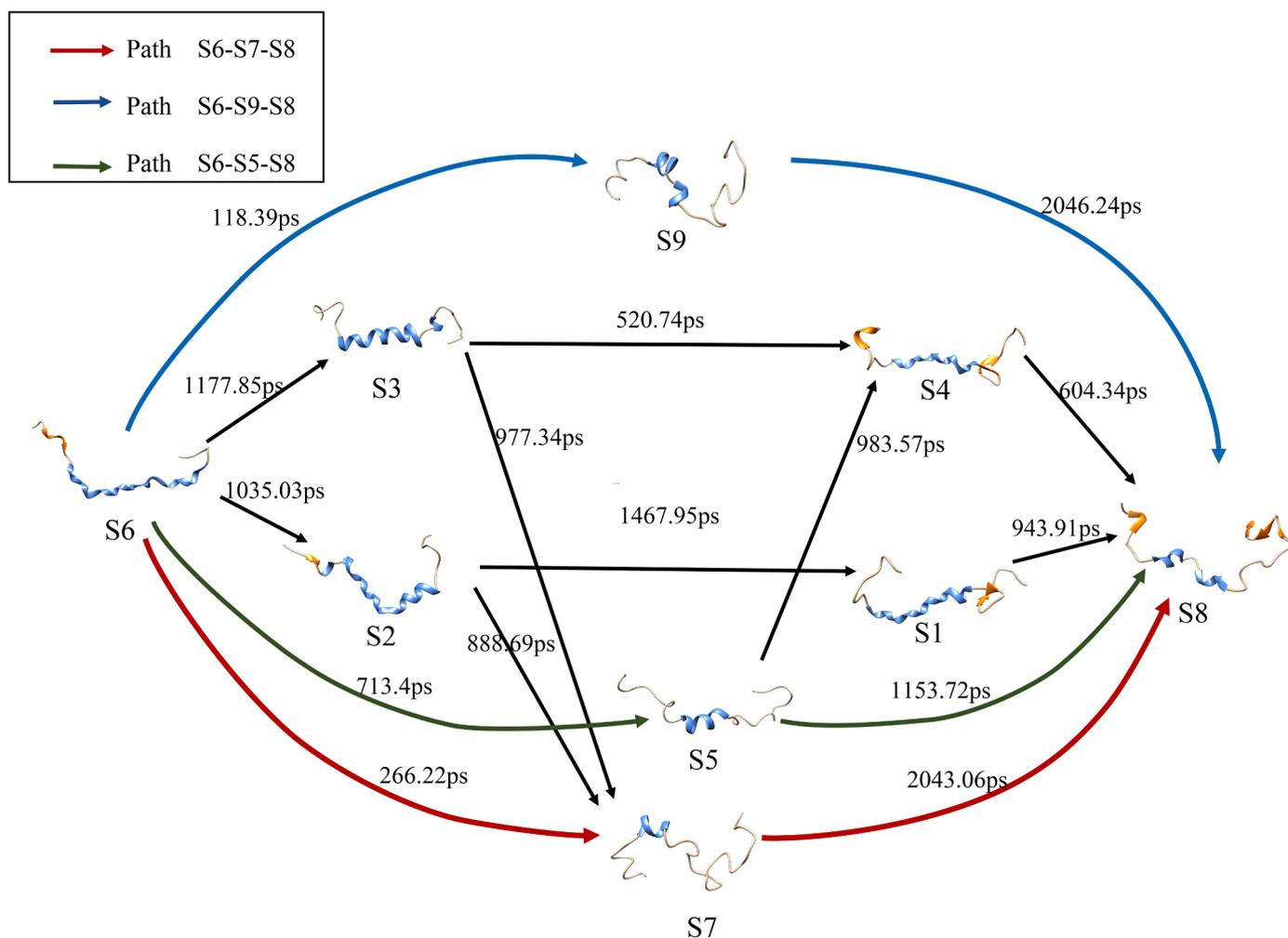
We further compare the contributions of each residue to PC1-PC3. As shown in Supplement Fig. 7, there are two regions that exhibit a higher contribution for three PCs. The first one is located within the Y10-A21 segment, and the other is in the V24-G37 segment. The results suggest that the primary distinction described by the first three PCs is the structural disparity between the two regions.

In order to determine whether there is a dynamic correlation between the structural fluctuations of two regions, we further assess the magnitude of the pairwise cross-correlation coefficients between any two residues of the A $\beta$  peptide. Supplement Fig. 8 is a dynamical cross-correlation map calculated using the 300 ns simulated trajectory. The two regions mentioned above, Y10-A21 and V24-G37, exhibit larger negative correlation values. This suggests that the structural fluctuation occurring in these two regions is in opposite directions. These results further confirm that the primary discrepancy among the sampled conformations is located in two specific regions. The detailed structural differences between A $\beta$ 42 peptide conformations will be discussed in the next section.

### 3.2.2. Free energy landscape of the sampled A $\beta$ peptide conformations

In order to clearly display the energy distribution of the sampled conformations, we calculated the relative free energy of each conformation using the probability density function. We then projected the relative energy onto the subspace defined by the first two eigenvectors of PCA. Fig. 7 displays the free energy landscape of all sampled conformations for 1IYT, 1Z0Q and APP. According to the free energy values, the energy landscape can be divided into four continuous regions, which are indicated by different colors in this figure. We further calculate the average secondary structure content of the conformations in different regions to differentiate between the conformations in those regions.

The conformations in the red-colored region have higher energy compared to the other three regions. Among these conformations, the proportion of coil structure is significantly higher than that of the other secondary structures, indicating that the A $\beta$  peptide is in a partially unfolded state. The conformations located in the orange-colored region near the red-colored region have lower energy than 8.9 kJ/mol. Based on the analysis of secondary structures, we can observe that the conformations of the A $\beta$  peptide in this region differ slightly from those in the red-colored region. The conformations with a relative energy of approximately 4.4 kJ/mol are grouped together in the green-colored region. Among these conformations, partial helical and sheet structures are found. The conformations with the lowest relative energy are located in the blue region. Comparing with the conformations in the other three regions, these structures have the highest sheet content and



**Fig. 8.** Top eight transition paths connecting from S6 to S8 and the corresponding MFPT between the two macro-states. The structure with the lowest free energy in each macro-state is displayed. Three paths, colored red, blue, and green respectively, are the major transition paths with the highest percentage contributions and the fewest macro-states.

the lowest coil content. This indicates that the A $\beta$  peptide has folded into stable conformations.

The energy landscape shows that conformations in different energy regions display distinct structural characteristics. A $\beta$  peptide conformations with higher energy are typically in a disordered state, characterized by a higher coil content. In contrast, the conformations located in the lower energy regions mainly adopt helix or sheet structures. The diverse conformations are consistent with the significant structural polymorphism observed in amyloid fibrils [41]. Different A $\beta$  peptide conformations may potentially lead to both on-pathway and off-pathway aggregation. Moreover, we speculate that these various conformations can serve as intermediates involved in the transition from helix to sheet. In order to clarify the process of this conformational transition, we attempted to construct a MSM model using these

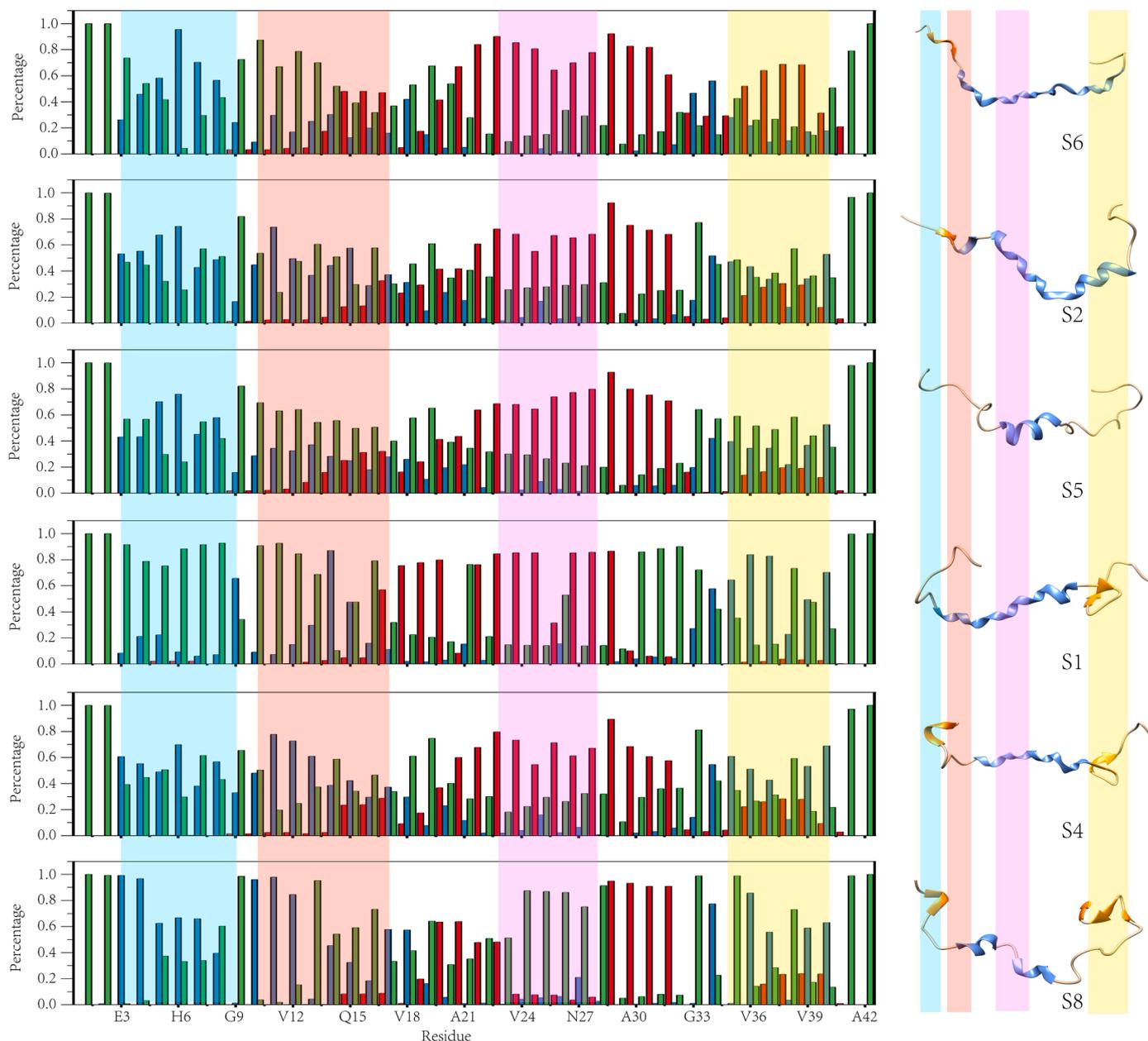
conformations.

### 3.3. The conformational transition described by MSM

#### 3.3.1. Data input and featurization

Although we have sampled the related A $\beta$  peptide conformations, it is difficult to obtain the possible transformation path and the detailed transition process solely by relying on the analysis of the simulated trajectories. Thus, we performed Markov state model (MSM) analysis on the multiple-rounds metadynamics trajectory to reveal the conformational transition pathway from helix to sheet.

As a powerful framework for reproducing the conformational dynamics of biomolecules using MD trajectory, the MSM model was applied to analyze the conformational transition process of the A $\beta$



**Fig. 9.** Distribution of secondary structure for each residue of the A $\beta$ 42 peptide along the S6-S2-S5-S1-S4-S8 pathway. For each residue of the A $\beta$ 42 peptide, the content of three secondary structures,  $\beta$ -sheet, helix, and coil, is calculated from 1000 conformations in each macro-state and colored blue, red and green, respectively. The regions involved in conformational changes are also highlighted in different colors. The average structure of each macro-state is also shown on the right side of the figure. Four primary changes are identified: a transition from intramolecular sheet to coil to sheet in the N-terminal region, a helix-to-coil transition between Y10 and L17, a collapse of the helical structure between D23 and K28, and a transition from helix to sheet in the C-terminal region.

peptide and subsequently identify possible aggregation pathways [42]. In order to capture the important conformational transitions, the raw metadynamics trajectory will first be transformed into a defined configuration space. This transformation will be based on the selected features that can accurately characterize the dynamic changes of the A $\beta$  peptide. In this study, we selected five features commonly used to describe conformational changes in a peptide. The VAMP2 score is calculated for five features at three different lag times. A higher VAMP2 score indicates a greater amount of information regarding the kinetic process of conformational transition described by this feature. The maximum dynamic variance can be achieved by utilizing this feature. Supplement Fig. 9 shows a comparison of VAMP2 scores among five features at three different lag times. While the VAMP2 score of backbone torsion angles and the selected Ca positions are somewhat similar at a lag time of 0.5 ns, the VAMP2 score of backbone torsion angle is significantly higher than that of the other four features at the other two lag times. Thus, the following analysis is performed using backbone torsion angle as the analytical feature.

### 3.3.2. Coordinate transform and discretization

After selecting the most appropriate feature, the raw metadynamics trajectory was transformed accordingly into the feature space defined by the selected feature. The transformed trajectory still has more degrees of freedom. Consequently, the TICA method is used to transform the feature space into a lower-dimensional space, which facilitates the discretization and analysis of the continuous trajectory.

By maximizing the autocorrelation of the data for a given lag time, tICA-projections of trajectories onto dominant independent components (ICs) can effectively extract the slow dynamical processes without losing any relevant information. In this study, we project the metadynamics trajectory onto the top four components and plot the marginal distributions of structures on these four components, as shown in Supplement Fig.10.A. Among the four ICs, the distribution of IC1 and IC3 overlapped to a lesser extent. The combined area of these two distributions accounts for the majority of the entire feature space. On the other hand, the distributions in the other two ICs mostly overlap. Moreover, the conformations of the A $\beta$  peptide projected onto the space defined by IC1 and IC3 are well positioned within the distinct regions displayed in Supplement Fig.10.B.

Then, all structures contained in the trajectory are clustered into a set of disjoint states using the k-means clustering method. The clustering is based on the backbone torsion angles in a subspace defined by IC1 and IC3. The continuous trajectory is transformed into a sequence of transitions, consisting of different states. The number of states is not only closely related to the accuracy of the transition sequence, but also heavily dependent on the computation time. By analyzing the change trend of the VAMP-2 score for different numbers of states, we can determine the optimal number of states. Supplement Fig. 11 shows the trend of VAMP-2 scores for state numbers ranging from 10 to 210. It has been found that the VAMP-2 scores at 40, 110, and 170 are approximately 3.95, which is significantly higher than the other values within the range of state numbers considered. Taking convenience into account, we set the state number to 40, as it has the lowest state number among the three states with the highest VAMP-2 score.

### 3.3.3. MSM Construction and validation

According to the above procedures, all conformations sampled in the metadynamics are grouped into 40 states. For a more concise representation that facilitates understanding of the process of conformational transition, the 40 states are further divided into several primary macro-states. These macro-states include multiple states that exhibit similar performance during transitions. It is important note that the process of conformational transition may be incomplete, even when using only a few macro-states. In this study, a total of 9 major macro-states were identified to construct the MSM model from a pool of 40 states. This was done based on the probability of each state given a macro-state and the

mutual transition between different states using the Perron Clustering Analysis method (PCCA+).

To verify the Markovianity of the resulting MSM model, a Chapman-Kolmogorov test is performed. Supplement Fig. 12 shows the results of the Chapman-Kolmogorov test. It can be seen that all predicted transition probabilities match very well with the estimated transition probabilities from one macro-state to the other eight macro-states. This confirms the Markovian nature of the model. The MSM model, which consists of nine macro-states, was used to analyze the process of conformational transition of the A $\beta$  peptide.

### 3.3.4. Transition among nine macro-states based on MSM model

The mean first pass time (MFPT) between two macro-states is calculated to infer the possible locations of macro-states in the transition process [43]. MFPT is a valuable kinetic property that describes conformational changes and quantifies the time it takes for a system to transition between two specific states. Moreover, MFPT also provides the contribution of both direct transitions between two states and indirect transitions that bypass intermediate states. Supplement Table 2 lists all MFPTs between two arbitrary macro-states. The range of MFPTs is from 28.1 to 2422.1 ps, with an average value of 870.2 ps for all 64 transitions. The smaller MFPT represents a higher transition probability and a smoother transition between two macro-states.

The summary of the inward and outward MFPT of each macro-state is displayed in Supplement Figure 13, along with their corresponding occupancy probabilities. We find that the probability of occupation varies from 4% to 34% among the nine macro-states. S1, S2, S3, and S4 have a small probability, less than 5%, which may be attributed to the larger inward MFPT and smaller outward MFPT. The lower probability of occupation means that these macro-states can frequently transform into other macro-states. On the contrary, S5, S6, S7, and S9 have a higher probability of being occupied due to their longer outward MFPT and shorter inward MFPT. It suggests that other macro-states can be easily transformed into these four specific macro-states. In addition, S8 has a relatively higher probability of 16% despite its inward MFPT being significantly larger than the outward MFPT. It implies that this state may represent the primary and stable conformation.

### 3.3.5. The transition paths from helix to sheet

In order to illustrate the structural variations among the nine macro-states, we calculated the average content of secondary structures based on a random selection of 1000 conformations from each macro-state. The distribution of content among coil, helix, and sheet is shown in Supplement Fig. 13.C. The coil is found to be the predominant structural form, with an average content exceeding 40% in all macro-states.

In this study, our aim is to describe the conformational transitions from helix to sheet. Thus, we intentionally select the macro-state with the highest helix and sheet content as the starting and ending state. As shown in Supplement Fig. 13. C, S6 has the highest helix content and can be considered the initial state. Although both S8 and S9 have the highest sheet content, we choose S8 as the terminal state because it has a higher inward MFPT compared to S9.

After defining the initial and terminal states, a total of 16 paths connecting these two states were found using transition path theory. The identified transition pathways are listed in Supplement Table 3, along with their individual and cumulative percentage contributions. Fig. 8 shows the top 8 transition pathways with an accumulated percentage contribution exceeding 90% and the representative structures for 9 macro-states.

Among the top three transition pathways, S5, S7, and S9 are the key intermediate states. Based on the comparison of the secondary structure distribution shown in Supplement Fig. 14, three pathways (S6-S5-S8, S6-S7-S8, S6-S9-S8) are associated with three primary structural changes. Firstly, the N-terminal portion of the A $\beta$  peptide undergoes a conformational transition from sheet to coil to sheet structures. Secondly, the helical structure in the CHC region shrinks, causing a disruption in the

interior helix from D23 to K28. Lastly, the helical structure in the C-terminal region transforms into a sheet structure.

The remaining 13 pathways were further analyzed to depict the conformational transitions in more detail and identify other key regions. By comparing the average MFPT values, we assessed the priority of these pathways. The smaller MFPT value indicates a higher probability of transitions between two states. Therefore, the pathway with the smallest average MFPT value will also be considered a possible transition pathway. The average MFPT values for the 13 pathways are listed in [Supplement Table 4](#). Although the pathways S6-S5-S4-S8 and S6-S3-S4-S8 have the lowest average of 767 ps, the pathway S6-S2-S5-S1-S4-S8 has a comparable MFPT to the above two pathways and contains the highest number of states. Therefore, this pathway is considered another transition pathway for the A $\beta$  peptide from helix to sheet.

According to the comparison of the secondary structure of the six states involved in this pathway, as shown in [Fig. 9](#), it can be observed that this pathway is associated with four conformational changes occurring in different regions of the A $\beta$  peptide. Firstly, a transition from an intramolecular sheet to a coil to an intramolecular sheet is observed in the N-terminal region of the A $\beta$  peptide. This finding is consistent with the results obtained from the three pathways: S6-S5-S8, S6-S7-S8, and S6-S9-S8. Secondly, a transition from a helix to a coil is observed between Y10 and L17. This conformational transition was also identified in our previous study [30].

The third structural change is the collapse of the helical structure that occurred in the CHC region of the A $\beta$  peptide. The transition from a helical structure to a coil structure is primarily observed in the middle of a long helix, specifically between residues D23 and K28. The corresponding MFPT value from helix (S4) to coil (S8) is 604.3 ps, which is smaller than the average value of 870.2 ps. Such structural changes are also found in S6-S5-S8, S6-S7-S8, and S6-S9-S8. At last, a transition from a helix to an intramolecular sheet is observed in the C-terminal region of the A $\beta$  peptide between M35 and V40. This transition is also found in S6-S5-S8, S6-S7-S8, and S6-S9-S8.

Based on the analysis of pathways obtained from the MSM model, we conclude that three fragile regions in the helical segment are more prone to transforming into other structural forms. These three regions may potentially be key sites associated with the transition from helix to sheet, and they may also play a regulatory role in the aggregation of A $\beta$  peptides.

Although the transition pathways were obtained from the conformational analysis of an individual peptide, the resulting pathways can be applicable to A $\beta$  monomers that adopt different conformations. This is because the sampled conformations cover almost all possible configurations of the single A $\beta$  peptide, as shown in the above analysis. The aggregation process, however, involves multiple A $\beta$  peptides, and the intermolecular interactions do play a mediating role in the conformational transition. Therefore, we will investigate this effect in future studies.

### 3.4. Verification of the fragile regions

Among the three fragile regions, Y10-L17 has been examined in the previous study, where we investigated the role played by this region in the aggregation of A $\beta$  peptides. D23-K28 and M35-V40 are two newly identified regions. Therefore, we further verify the structural changes and potential roles in the aggregation of A $\beta$  peptides through hydrogen bond analysis and by examining the binding affinity of the aggregation inhibitors.

#### 3.4.1. Analysis of hydrogen bonds

The network of hydrogen bonds is important for maintaining the helical structure of a peptide. In this study, we analyzed the hydrogen bonds within the fragment spanning from F20 to A42. [Supplement Figure 15](#) plots the probability of hydrogen bonds calculated from nine macro-states, respectively. We can observe that residues D23-K28 and

M35-V40 exhibit a similar trend in hydrogen bond probability. The higher probability is found for S1-S3 and S5, as well as S6, indicating that the helical structure is formed in these intermediate states in two regions. On the contrary, the probability is decreased for S4, S7, S8, and S9. The analysis of hydrogen bonds further demonstrates that the helical structure in D23-K28 and M35-V40 is changes into other structural forms.

#### 3.4.2. Molecular docking of the aggregation inhibitors

In the above analysis, we identified two regions where the helical structure is more likely to transform into a coil or an intramolecular sheet structure. It has been accepted that the formation of an intramolecular sheet structure is crucial for the subsequent aggregation of A $\beta$  peptides. Therefore, these two regions, D23-K28 and M35-V40, may possibly play a role in the aggregation of A $\beta$  peptides. In order to verify whether these two regions are responsible for aggregation, we conducted a docking study to explore the binding affinity of these two regions with inhibitors of amyloid protein aggregation. A total of 735 known aggregation inhibitors, along with their inhibitory activity, were obtained from the PubChem database (<https://pub-chem.ncbi.nlm.nih.gov/>). Each inhibitor was docked with nine macro-states using Vina software [44]. The docking box is placed to cover these two areas. The docking results of nine intermediate states are shown in [Fig. 10](#). In order to clearly demonstrate the docking profile of these aggregation inhibitors, we selected a binding energy threshold of  $-5$  kcal/mol. This threshold was also utilized in another study [45].

As shown in [Fig. 10](#), we found that the binding energy of approximately 90% of the aggregation inhibitors is lower than the threshold value. Most inhibitors have a higher affinity for two specific regions, indicating that these two key regions may be involved in the aggregation of the A $\beta$  peptide. The docking results indirectly confirm the significance of D23-K28 and M35-V40 in the aggregation of A $\beta$  peptides.

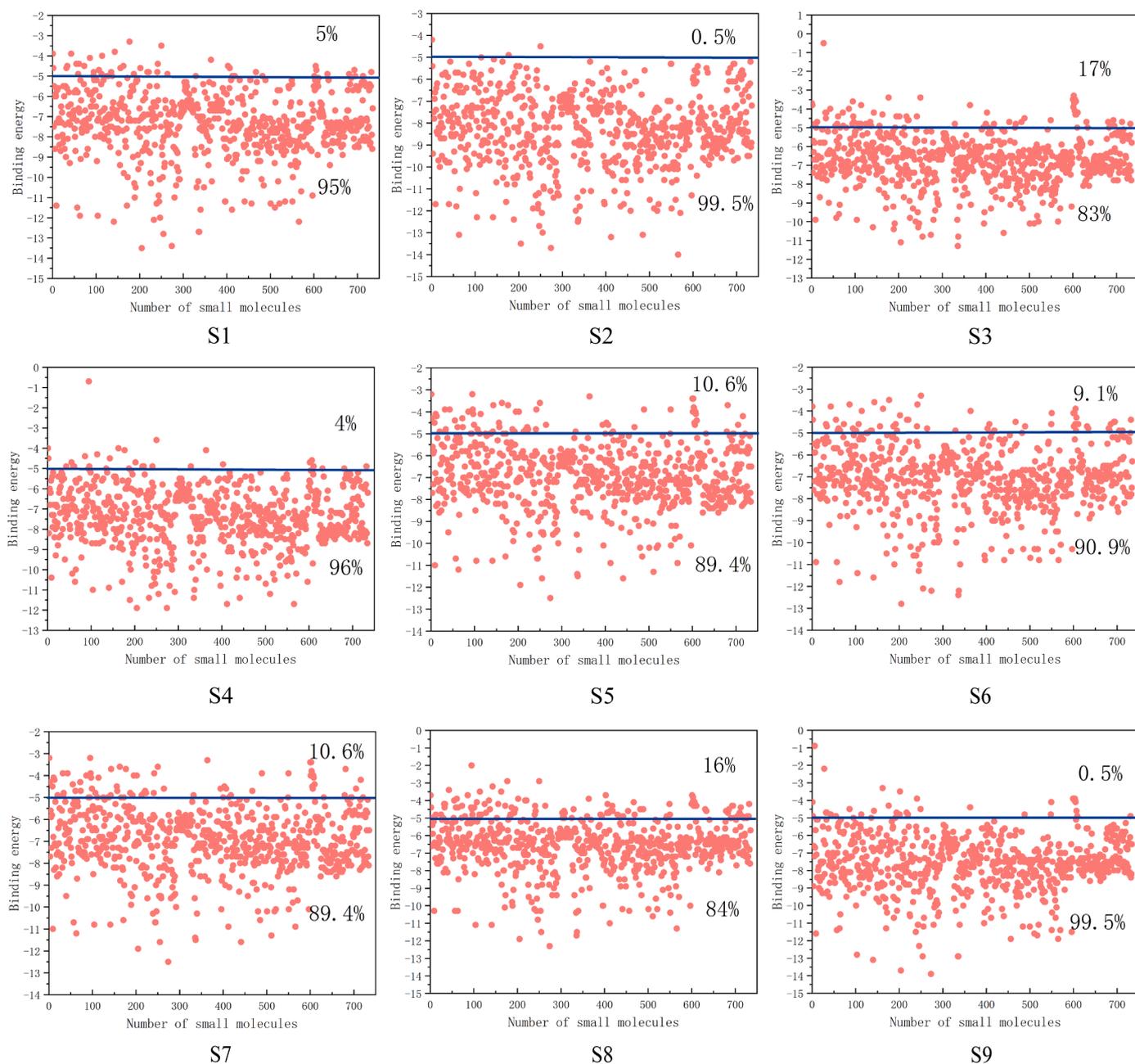
## 4. Conclusion

It is well known that a high content of intramolecular  $\beta$ -sheet structure is important for the aggregation of A $\beta$  peptide and the formation of A $\beta$  fibers. But after being cleaved from APP, the A $\beta$  peptide still adopts an  $\alpha$ -helical conformation. Therefore, the conformational transition from helix to sheet is a crucial step in the aggregation of the A $\beta$  peptide. However, the structural details of this process still remain unclear due to the heterogeneity and transient nature of the A $\beta$  peptide. In this study, we employ an enhanced sampling strategy that combines ANN with metadynamics to investigate the conformational transition from a helix to a sheet structure in the A $\beta$ 42 peptide. The strategy consists of two parts: applying ANN to optimize CVs and conducting metadynamics based on the optimized CVs to sample conformations. Compared with ordinary MD and the previous Hamiltonian Exchange MD, this strategy achieves better sampling performance in terms of the distribution of sampled conformations. The sampling efficiency is increased by 10-fold compared to the Hamilton Exchange MD and by 1000-fold compared to the ordinary MD.

In addition, we constructed a Markov state model using the sampled conformations of the A $\beta$ 42 peptide. The transition process from helix to sheet was described, and three key regions closely associated with this transition were identified. The collapse of the long helix between D23-K28 and the transition from helix to sheet in M35-V40 are observed for the first time, to the best of our knowledge. The docking results indicate that most of the inhibitors of amyloid aggregation can bind well to these two regions. These two regions can be considered potential targets for screening and designing inhibitors of A $\beta$  peptide aggregation.

### CRedit authorship contribution statement

Huilin Wen : Main works including extraction and optimization of CVs and metadynamics and MSM, Writing Original Draft. Hao Ouyang:



**Fig. 10.** Binding energy profiles of 735 aggregation inhibitors across 9 macro states. For each subgraph, a line is drawn to represent the binding energy threshold of  $-5$  kcal/mol. The ratios above and below the energy threshold are displayed on both sides of the threshold line. We observe that the binding energy of more than 90% of the inhibitors is lower than the threshold value in all macro-states, except for S3 and S8, where the ratio is below 85%.

Verification using virtual screening. Hao Shang: Data curation and visualization. Chaohong Da: Verification using virtual screening. Tao Zhang (**Corresponding Author**): Writing, review and editing manuscript. All authors have read and approved the final manuscript.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgement

This research was funded by National Natural Science Foundation of China, grant number 31300597.

#### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.csbj.2023.12.015](https://doi.org/10.1016/j.csbj.2023.12.015).

#### References

- [1] 2023 Alzheimer's disease facts and figures. *Alzheimers Dement* 2023;19(4): 1598–695.
- [2] Ferrari C, Sorbi S. The complexity of Alzheimer's disease: an evolving puzzle. *Physiol Rev* 2021;101(3):1047–81.
- [3] Knopman DS, Amieva H, Petersen RC, et al. Alzheimer disease. *Nat Rev Dis Prim* 2021;7(1):33.
- [4] Hampel H, Vassar R, De Strooper B, et al. The beta-secretase BACE1 in Alzheimer's disease. *Biol Psychiatry* 2021;89(8):745–56.
- [5] Miyashita N, Straub JE, Thirumalai D. Structures of beta-amyloid peptide 1-40, 1-42, and 1-55-the 672-726 fragment of APP-in a membrane environment with

- implications for interactions with gamma-secretase. *J Am Chem Soc* 2009;131(49):17843–52.
- [6] Meng F, Bellaiche MMJ, Kim JY, et al. Highly disordered amyloid-beta monomer probed by single-molecule FRET and MD simulation. *Biophys J* 2018;114(4):870–84.
- [7] Löhr T, Kohlhoff K, Heller GT, et al. A kinetic ensemble of the Alzheimer's A $\beta$  peptide. *Nat Comput Sci* 2021;1(1):71–8.
- [8] Chimon S, Shaibat MA, Jones CR, et al. Evidence of fibril-like beta-sheet structures in a neurotoxic amyloid intermediate of Alzheimer's beta-amyloid. *Nat Struct Mol Biol* 2007;14(12):1157–64.
- [9] Khaled M, Ronnback I, Ilag LL, et al. A hairpin motif in the amyloid-beta peptide is important for formation of disease-related oligomers. *J Am Chem Soc* 2023;145(33):18340–54.
- [10] Aleksis R, Oleskovs F, Jaudzems K, et al. Structural studies of amyloid-beta peptides: unlocking the mechanism of aggregation and the associated toxicity. *Biochimie* 2017;140:176–92.
- [11] Abelein A, Abrahams JP, Danielsson J, et al. The hairpin conformation of the amyloid beta peptide is an important structural motif along the aggregation pathway. *J Biol Inorg Chem* 2014;19(4-5):623–34.
- [12] Itoh SG, Okumura H. Hamiltonian replica-permutation method and its applications to an alanine dipeptide and amyloid-beta(29-42) peptides. *J Comput Chem* 2013;34(29):2493–7.
- [13] Itoh SG, Okumura H. Oligomer formation of amyloid-beta(29-42) from its monomers using the hamiltonian replica-permutation molecular dynamics simulation. *J Phys Chem B* 2016;120(27):6555–61.
- [14] Itoh SG, Yagi-Utsumi M, Kato K, et al. Key residue for aggregation of amyloid-beta peptides. *ACS Chem Neurosci* 2022;13(22):3139–51.
- [15] Kirkitadze MD, Condrum MM, Teplow DB. Identification and characterization of key kinetic intermediates in amyloid beta-protein fibrillogenesis. *J Mol Biol* 2001;312(5):1103–19.
- [16] Misra P, Kodali R, Chemuru S, et al. Rapid alpha-oligomer formation mediated by the Abeta C terminus initiates an amyloid assembly pathway. *Nat Commun* 2016;7:12419.
- [17] Pagano K, Tomaselli S, Molinari H, et al. Natural compounds as inhibitors of abeta peptide aggregation: chemical requirements and molecular mechanisms. *Front Neurosci* 2020;14:619667.
- [18] Iadanza MG, Jackson MP, Hewitt EW, et al. A new era for understanding amyloid structures and disease. *Nat Rev Mol Cell Biol* 2018;19(12):755–73.
- [19] Saravanan KM, Zhang H, Zhang H, et al. On the conformational dynamics of beta-amyloid forming peptides: a computational perspective. *Front. Bioeng Biotechnol* 2020;8:532.
- [20] Kotler SA, Brender JR, Vivekanandan S, et al. High-resolution NMR characterization of low abundance oligomers of amyloid-beta without purification. *Sci Rep* 2015;5:11811.
- [21] Jiang D, Rauda I, Han S, et al. Aggregation pathways of the amyloid beta(1-42) peptide depend on its colloidal stability and ordered beta-sheet stacking. *Langmuir* 2012;28(35):12711–21.
- [22] Xie H, Rojas A, Maisuradze GG, et al. Mechanistic kinetic model reveals how amyloidogenic hydrophobic patches facilitate the amyloid-beta fibril elongation. *ACS Chem Neurosci* 2022;13(7):987–1001.
- [23] K.D. Nadezhdin OVB, Bocharov EV, Arseniev AS. Structural and dynamic study of the transmembrane domain of the amyloid precursor protein. *Acta Nat* 2011;3:69–76.
- [24] Chen W, Gamache E, Rosenman DJ, et al. Familial Alzheimer's mutations within APPTM increase Abeta42 production by enhancing accessibility of epsilon-cleavage site. *Nat Commun* 2014;5:3037.
- [25] Takeda T, Klimov DK. Interpeptide interactions induce helix to strand structural transition in Abeta peptides. *Proteins* 2009;77(1):1–13.
- [26] Tomaselli S, Esposito V, Vangone P, et al. The alpha-to-beta conformational transition of Alzheimer's Abeta-(1-42) peptide in aqueous media is reversible: a step by step conformational analysis suggests the location of beta conformation seeding. *Chembiochem* 2006;7(2):257–67.
- [27] Mudedla SK, Murugan NA, Agren H. Effect of familial mutations on the interconversion of alpha-helix to beta-sheet structures in an amyloid-forming peptide: insight from umbrella sampling simulations. *ACS Chem Neurosci* 2019;10(3):1347–54.
- [28] Saini RK, Shuaib S, Goyal D, et al. Impact of mutations on the conformational transition from alpha-helix to beta-sheet structures in arctic-type abeta(40): insights from molecular dynamics simulations. *ACS Omega* 2020;5(36):23219–28.
- [29] Strodel B. Amyloid aggregation simulations: challenges, advances and perspectives. *Curr Opin Struct Biol* 2021;67:145–52.
- [30] Chakraborty D, Straub JE, Thirumalai D. Energy landscapes of Abeta monomers are sculpted in accordance with Ostwald's rule of stages. *Sci Adv* 2023;9(12):eadd6921.
- [31] Chakraborty D, Straub JE, Thirumalai D. Differences in the free energies between the excited states of Abeta40 and Abeta42 monomers encode their aggregation propensities. *Proc Natl Acad Sci USA* 2020;117(33):19926–37.
- [32] Paul A, Samantray S, Anteghini M, et al. Thermodynamics and kinetics of the amyloid-beta peptide revealed by Markov state models based on MD data in agreement with experiment. *Chem Sci* 2021;12(19):6652–69.
- [33] Qiu T, Liu Q, Chen YX, et al. Abeta42 and Abeta40: similarities and differences. *J Pept Sci* 2015;21(7):522–9.
- [34] Meisl G, Yang X, Hellstrand E, et al. Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Abeta40 and Abeta42 peptides. *Proc Natl Acad Sci USA* 2014;111(26):9384–9.
- [35] Tiwary P, Parrinello M. From metadynamics to dynamics. *Phys Rev Lett* 2013;111(23):230602.
- [36] Trapl D, Horvancanin I, Mareska V, et al. Anncolvar: approximation of complex collective variables by artificial neural networks for analysis and biasing of molecular simulations. *Front Mol Biosci* 2019;6:25.
- [37] Zhou M, Wen H, Lei H, et al. Molecular dynamics study of conformation transition from helix to sheet of Abeta42 peptide. *J Mol Graph Model* 2021;109:108027.
- [38] Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015;1-2:19–25.
- [39] Husic BE, Pande VS. Markov state models: from an art to a science. *J Am Chem Soc* 2018;140(7):2386–96.
- [40] Scherer MK, Trendelkamp-Schroer B, Paul F, et al. PyEMMA 2: a software package for estimation, validation, and analysis of markov models. *J Chem Theory Comput* 2015;11(11):5525–42.
- [41] Lutter L, Aubrey LD, Xue WF. On the structural diversity and individuality of polymorphic amyloid protein assemblies. *J Mol Biol* 2021;433(20):167124.
- [42] Illig AM, Strodel B. Performance of markov state models and transition networks on characterizing amyloid aggregation pathways from MD data. *J Chem Theory Comput* 2020;16(12):7825–39.
- [43] Polizzi NF, Therien MJ, Beratan DN. Mean first-passage times in biology. *Isr J Chem* 2016;56(9-10):816–24.
- [44] Eberhardt J, Santos-Martins D, Tillack AF, et al. AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J Chem Inf Model* 2021;61(8):3891–8.
- [45] Xia J, Hu JN, Wang Z, et al. Based on network pharmacology and molecular docking to explore the protective effect of Epimedii Folium extract on cisplatin-induced intestinal injury in mice. *Front Pharm* 2022;13:1040504.