

# The Monomeric Conformational Ensembles of A $\beta$ 40 and A $\beta$ 42 Encode Their Differential Amyloid Aggregation Propensity

Irene Cadenelli, Andrea Ciccolo, Andrea Tagliabue, Giulia Rossi, Valeria Conti Nibali,\* and Davide Bochicchio\*



Cite This: *J. Phys. Chem. B* 2026, 130, 6035–6046



Read Online

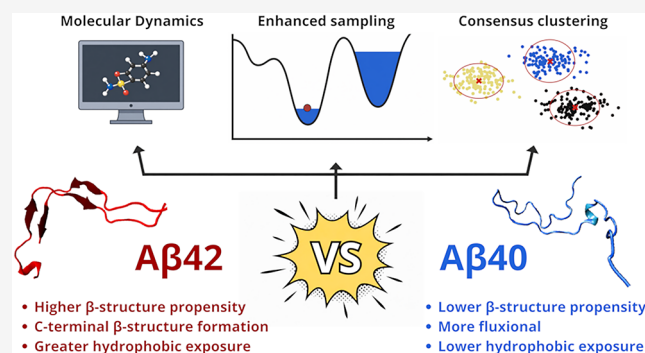
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** A $\beta$ 40 and A $\beta$ 42 peptides differ by just two C-terminal residues, yet they display strikingly different aggregation and toxicity profiles. Whether this distinction is already encoded at the monomer level is still under debate. Here, we combine extensive all-atom simulations in explicit solvent, well-tempered metadynamics, and a tailored consensus cluster analysis to compare the monomeric ensembles of the two isoforms under identical conditions. Both peptides populate broad, coil-like conformational distributions; however, A $\beta$ 42 shows a systematically higher  $\beta$ -structure propensity, especially in the C-terminal region, and samples more extended conformations with higher hydrophobic exposure compared to A $\beta$ 40. These results support a mechanistic link between sequence-encoded monomer conformational preferences and the differential amyloidogenicity of the two isoforms, highlighting monomer-level determinants of A $\beta$ 42's distinct aggregation behavior.



highlighting monomer-level determinants of A $\beta$ 42's distinct

## INTRODUCTION

Alzheimer's Disease (AD) is a widespread neurodegenerative disorder characterized by a progressive decline of cognitive function, dysfunction, and neuronal loss,<sup>1–4</sup> and currently ranks among the leading causes of death and disability worldwide, particularly in older populations.<sup>5</sup> The main hallmarks of AD are the extracellular deposits of amyloid fibrils formed by amyloid- $\beta$  (A $\beta$ ) peptides, predominantly by the A $\beta$ 40 and A $\beta$ 42 isoforms. Although these two isoforms differ by only two C-terminal residues, they display markedly distinct aggregation kinetics, morphologies, and toxicities.<sup>6,7</sup> Compared with A $\beta$ 40, A $\beta$ 42 is more abundant in mature amyloid fibrils, even though A $\beta$ 40 is produced in higher amounts from the amyloid- $\beta$  precursor protein.<sup>8</sup> In addition, A $\beta$ 42 aggregates more rapidly and is more capable of damaging neuronal membranes by aggregating into pore-forming oligomers.<sup>6,7,9</sup>

Amyloid oligomers (4–10 monomers) can be significantly more toxic than micrometer-long, mature fibrils.<sup>10,11</sup> Fibril-catalyzed *secondary nucleation*—the formation of new nuclei on the surface of existing fibrils—can dominate aggregate proliferation and enhance the generation of oligomeric toxic species, especially for A $\beta$ 42.<sup>6,7,12,13</sup>

At the monomer level, both A $\beta$ 40 and A $\beta$ 42 are intrinsically disordered peptides (IDPs) that populate very broad conformational ensembles in aqueous solution.<sup>14</sup> Single-molecule spectroscopy and all-atom simulations indicate highly

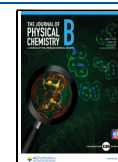
dynamic, coil-like distributions with rapid interconversion among conformers.<sup>15</sup> The vast heterogeneity of possible structures at the monomer level is reflected in a certain degree of polymorphism of amyloid fibrils. However, the number of structures identified for amyloid fibrils is substantially lower than the number of monomeric conformers. Therefore, whether and how subtly sequence-encoded monomer features can bias downstream aggregation pathways into different structures remains a key open problem.<sup>16,17</sup> A recent work demonstrated that, despite amyloid polymorphism, the sequence regions most often forming the structural core of amyloid fibrils are those exhibiting high intrinsic  $\beta$ -sheet propensity in the monomer.<sup>18</sup> This observation suggests that, in principle, the aggregation propensity of different isoforms could be anticipated from their distinct monomeric conformational preferences in solution.<sup>19</sup> Experimentally, however, isolating and characterizing truly monomeric A $\beta$  species remains extremely challenging, as IDPs rapidly self-associate even at submicromolar concentrations, populate fast mono-

**Received:** February 13, 2026

**Revised:** June 3, 2026

**Accepted:** June 3, 2026

**Published:** June 9, 2026



mer–oligomer equilibria, and quickly adsorb onto experimental substrates.<sup>20</sup>

In this context, molecular simulations with atomistic resolution have become essential tools for understanding these complex systems. Methodological developments now enable a more reliable characterization of the disordered states of proteins. In particular, protein force fields refined for IDPs (e.g., CHARMM36m, a99SB-disp) have substantially improved agreement with SAXS, NMR, and smFRET benchmarks for unfolded and disordered proteins.<sup>21–23</sup> In parallel, enhanced-sampling strategies, such as well-tempered metadynamics, combined with time-independent reweighting, allow reconstruction of unbiased thermodynamics over slow collective variables (CVs), providing statistically robust estimates of conformational populations and free energy basins for flexible peptides.<sup>24–26</sup>

Furthermore, machine-learning approaches have begun to enhance our understanding of the conformational landscapes and transition pathways of amyloid  $\beta$  peptides in solution. Notable examples include neural-network Markov models, such as VAMPNet/CoVAMPnet,<sup>27,28</sup> which learn soft state decompositions and kinetics, as well as graph-based variants that scale and regularize these analyses on large data sets,<sup>29</sup> and deep generative models that sample IDP-consistent conformers.<sup>30–32</sup>

Prior simulation work on  $A\beta$  peptides showed clearly that, under physiological conditions,  $A\beta 40$  and  $A\beta 42$  populate an ensemble of random-coil–like conformations, even if secondary structures along the sequence are transiently sampled.<sup>33–35</sup> Bias-Exchange Molecular Dynamics simulations guided by NMR data have revealed an inverted free energy landscape: the global minimum corresponds to an ensemble of highly disordered conformations, whereas regions of higher free energy are associated with more structured states displaying an increasing content of secondary structure elements.<sup>36</sup> For both isoforms, the  $\alpha$ -helix content is generally lower than the  $\beta$ -sheet content, consistent with circular dichroism data indicating an  $\alpha$ -helix content of  $\approx 9\%$  and a  $\beta$ -content of 12–25%.<sup>10,37</sup> However, it has also been suggested that there are isoform-specific differences, especially in the balance between compact and extended conformations, as well as in the organization of hydrophobic contacts and C-terminal  $\beta$ -propensities.<sup>38</sup> Still, consensus has been limited by well-known force field and water-model biases affecting the dimensions and secondary-structure propensities of intrinsically disordered proteins,<sup>22,23,39</sup> as well as by the lack of consistency in force fields, simulation protocols, and analysis pipelines across computational studies of  $A\beta 40$  and  $A\beta 42$  in solution.<sup>40–42</sup>

In this study, we adopt a systematic computational strategy to determine whether structural differences between  $A\beta 42$  and  $A\beta 40$  already emerge at the monomer level. Using extensive all-atom simulations in explicit solvent with an IDP-calibrated force field, combined with advanced enhanced sampling, rigorous reweighting, and a clustering framework that robustly identifies dominant conformational families across statistically equivalent trajectory subsets, we compare the monomeric ensembles of the two isoforms under identical conditions. This approach allows us to resolve isoform-specific differences in secondary structure propensities, hydrophobic exposure, and contact patterns that are consistent with known fibril-forming interfaces.<sup>18,43,44</sup> Our results suggest that sequence-encoded

monomeric conformational preferences contribute to the distinct aggregation behavior of  $A\beta 40$  and  $A\beta 42$ .

## METHODS

### System Setup and Molecular Dynamics Simulations

All simulations of  $A\beta 40$  and  $A\beta 42$  were performed using identical protocols. The peptide sequences were taken from the Protein Data Bank ( $A\beta 42$ : PDB ID 6SZF;  $A\beta 40$ : PDB ID 1BA4). In both peptides, the C-terminal residue (A42 in  $A\beta 42$ ; V40 in  $A\beta 40$ ) was modeled with a deprotonated carboxylate group. The N-terminal aspartic acid carried a protonated amino group whose charge was neutralized by the negatively charged side chain, resulting in an overall neutral residue. Side-chain protonation states were assigned assuming physiological pH: lysine and arginine were kept protonated; nonterminal aspartic acid and glutamic acid were modeled in their deprotonated forms; and histidine was assigned as the neutral tautomer, which is the predominant state near-physiological pH.<sup>45</sup>

Each peptide was solvated in a cubic box of TIP3P water with an edge length of approximately 10 nm, under periodic boundary conditions in all directions. The box size was chosen to ensure at least 1.2–1.5 nm of solvent between the peptide and its periodic images, thereby excluding spurious interimage interactions. Na<sup>+</sup> and Cl<sup>−</sup> ions were added to neutralize the system and reach a NaCl concentration of 0.15 M. All simulations employed the CHARMM36m force field,<sup>21</sup> which provides an improved balance between secondary-structure propensities and intrinsic disorder relative to earlier CHARMM versions.<sup>21,39,46</sup>

All simulations were carried out using GROMACS<sup>47</sup> patched with PLUMED<sup>48</sup> for enhanced sampling. Production trajectories were generated in the NPT ensemble using a 2 fs time step. For each peptide, three independent replicas of 3  $\mu$ s were performed, yielding a total simulation time of 9  $\mu$ s per system. Coordinates were saved every 20 ps. The temperature was maintained at 310 K using the V-rescale thermostat,<sup>49</sup> and pressure was maintained at 1 atm using the isotropic Parrinello–Rahman barostat.<sup>50,51</sup>

Electrostatic interactions were computed using Particle Mesh Ewald. Lennard-Jones interactions were treated using a force-switch scheme consistent with CHARMM36m: the van der Waals potential was smoothly switched between 1.0 and 1.2 nm, and the real-space cutoff for Coulomb interactions was set at 1.2 nm. Long-range dispersion corrections were disabled as required for simulations using the CHARMM force field. Unless otherwise specified, simulations employed the default parameters of GROMACS (ver. 2024.3) and PLUMED (ver. 2.9.3).

### Collective Variables and Metadynamics Simulations

**Collective Variables.** Enhanced sampling was performed using two collective variables (CVs) describing secondary structure content: the  $\alpha$ - and  $\beta$ -content, implemented in PLUMED as ALPHARMSD and ANTIBETARMSD.<sup>52</sup> These CVs quantify the amount of  $\alpha$ -helical and antiparallel  $\beta$ -sheet structure by comparing contiguous residue segments to ideal  $\alpha$  and  $\beta$  blocks derived from a statistical analysis of protein structures in the PDB. Specifically:

- The  $\alpha$ -content is obtained by evaluating all contiguous 6-residue segments and comparing them to the ideal  $\alpha$ -helical block.
- The  $\beta$ -content is obtained by evaluating all contiguous 8-residue segments (with a two-residue separation between the central triplets) against the ideal antiparallel  $\beta$ -sheet block.

The similarity is quantified via the switching function

$$S = \sum_i \frac{1 - \left(\frac{r_i - d_0}{r_0}\right)^n}{1 - \left(\frac{r_i - d_0}{r_0}\right)^m} \quad (1)$$

where  $r_i$  is the RMSD of each segment. Parameters were set to  $d_0 = 0$ ,  $r_0 = 0.1$ ,  $n = 8$ , and  $m = 12$ , as prescribed in ref 52.

Results are presented in terms of the CVs  $n_\alpha$  and  $n_\beta$ , obtained by rescaling the raw  $\alpha$ - and  $\beta$ -content CVs by factors of 1.1 and 2.5, respectively. These rescaled quantities provide an approximate estimate of the number of residues engaged in  $\alpha$ -helical or  $\beta$ -like structures, consistent with typical definitions used in secondary-structure assignment tools such as STRIDE<sup>53</sup> and DSSP.<sup>54</sup> The optimization of the rescaling factors and their validation against STRIDE and DSSP assignments are described in detail in the SI.

**Metadynamics.** Enhanced sampling was performed using well-tempered metadynamics (MetaD).<sup>25</sup> Gaussian biases were deposited every 5 ps, with an initial height of 0.5 kJ/mol and an initial width of 0.1 (in CV units). The Gaussian width was chosen based on the natural fluctuations of the  $\alpha$ - and  $\beta$ -content CVs, ensuring efficient filling of free energy basins without oversmoothing the landscape. For computational efficiency, a grid in CV space was used to store the bias potential, with a resolution of 0.01. A bias factor of  $\gamma = 10$  was selected to enhance sampling while restricting the exploration to conformations with physically reasonable free energies, thereby avoiding systematic visitation of highly unfavorable, nonrepresentative states.

### Free Energy Estimation

Free energy surfaces (FESs) were reconstructed from the MetaD bias using the well-tempered metadynamics (WT-MetaD) formalism,<sup>25</sup> as implemented in the `sum_hills` utility of PLUMED.<sup>48</sup> In the well-tempered scheme, the time-dependent bias  $V(\xi, t)$  converges to a rescaled estimate of the underlying free energy

$$\lim_{t \rightarrow \infty} V(\xi, t) = -\frac{\gamma - 1}{\gamma} F(\xi) + C$$

where  $F(\xi)$  is the physical free energy as a function of the collective variable  $\xi$  and  $C$  is an irrelevant additive constant. The bias is constructed as

$$V(\xi, t) = \sum_{n\tau < t} h(n\tau) \exp\left[-\frac{(\xi - \xi(n\tau))^2}{2\sigma^2}\right]$$

where  $\sigma$  is the Gaussian width and  $h(n\tau)$  is the time-dependent height

$$h(n\tau) = h_0 \exp\left[-\frac{\beta V(\xi(n\tau), n\tau)}{\gamma - 1}\right]$$

with  $h_0$  the initial Gaussian height and  $\beta = 1/(k_B T)$ .

Throughout the manuscript, FESs are expressed in terms of the structural variables  $n_\alpha$  and  $n_\beta$  (see [Collective Variables and Metadynamics Simulations](#) Section), which provide approximate estimates of the number of residues participating in  $\alpha$ -helical or  $\beta$ -like structures.

**Convergence Assessment.** Convergence was assessed by monitoring (i) the time evolution of the accumulated MetaD bias and (ii) the temporal stability of the relative populations associated with the dominant free energy minima. Convergence metrics reported in the [Supporting Information](#) show that bias deposition reaches a quasi-stationary regime and that the main FES features remain stable over the final portion of each trajectory, supporting the qualitative and comparative analyses presented for A $\beta$ 40 and A $\beta$ 42.

### Reweighting Procedure

Because WT-MetaD modifies the sampling distribution, unbiased ensemble averages were recovered by reweighting the biased trajectories. We used the standard final-bias approximation for WT-MetaD, which treats the bias accumulated at the end of the simulation as a quasi-static bias and is therefore formally analogous to static-bias reweighting, as in umbrella sampling.<sup>55</sup> Using the notation introduced above, with  $\xi = (n_\alpha, n_\beta)$ , the statistical weight assigned to configuration  $i$  was written as

$$w_i \propto \exp[\beta V(\xi_i, t_{\text{final}})] \quad (2)$$

where  $V(\xi, t_{\text{final}})$  is the final WT-MetaD bias added to the system Hamiltonian. In the quasi-stationary regime, the WT-MetaD bias is related to the underlying free energy surface by

$$V(\xi, t_{\text{final}}) \simeq -\left(1 - \frac{1}{\gamma}\right) F(\xi) + C \quad (3)$$

where  $\gamma$  is the bias factor.<sup>25,56</sup> Therefore, the weights can equivalently be expressed as

$$\log w_i = -\beta \left(1 - \frac{1}{\gamma}\right) F(n_{\alpha,i}, n_{\beta,i}) + C' \quad (4)$$

which is the form used in our implementation.

To ensure that the final-bias approximation was applied only under quasi-stationary conditions, the initial 1  $\mu$ s of each trajectory, during which the bias deposition had not yet stabilized, was excluded from all reweighted analyses. The two-dimensional FES  $F(n_\alpha, n_\beta)$  was reconstructed from the combined sampling of the three independent replicas, and  $F(n_{\alpha,i}, n_{\beta,i})$  was obtained for each configuration by bilinear interpolation over the replica-averaged FES grid.

To avoid numerical underflow, we applied a standard log-shift,

$$\log w_i^{(\text{shift})} = \log w_i - \max_j \log w_j \quad (5)$$

which preserves relative weights and does not affect normalized ensemble averages. Reweighted histograms were then computed as

$$H_{\text{rw}}(x) = \sum_i w_i^{(\text{shift})} \delta(x - x_i) \quad (6)$$

The final unbiased distribution was obtained by averaging the independently reweighted per-replica histograms,

$$\bar{H}(x) = \frac{1}{N_{\text{rep}}} \sum_{r=1}^{N_{\text{rep}}} H_{\text{rw}}^{(r)}(x) \quad (7)$$

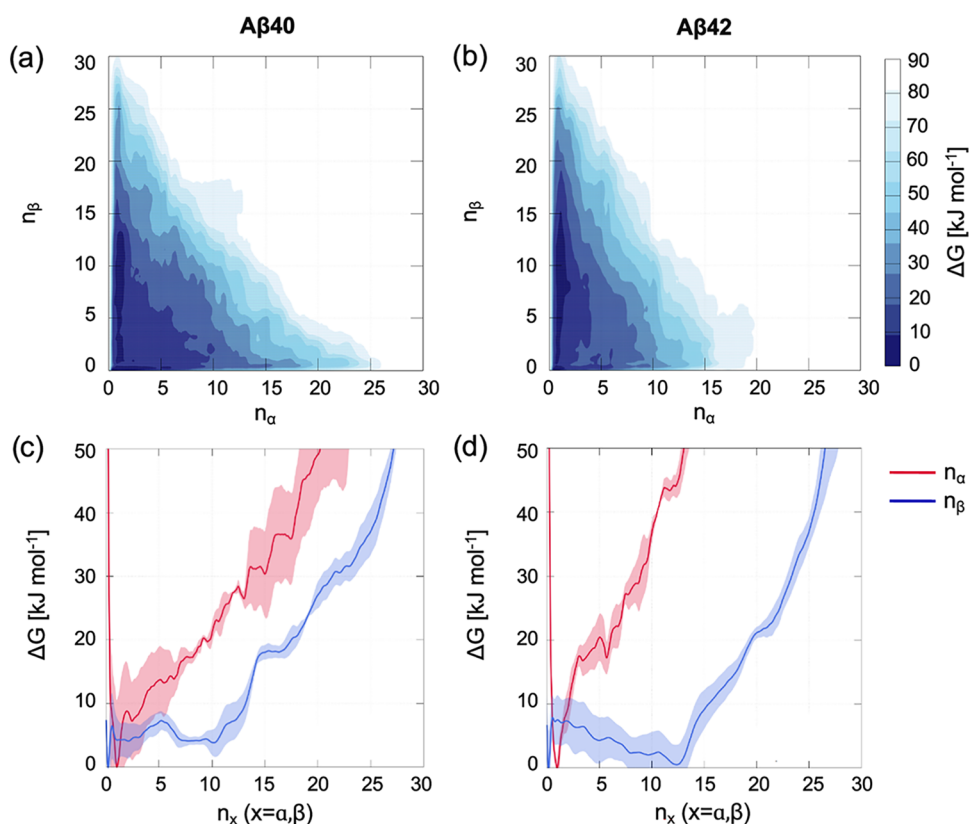
and uncertainties were reported as the standard error of the mean across replicas. This procedure follows the theoretical framework of WT-MetaD and its associated reweighting theory.<sup>25,26,56</sup> Alternative time-independent approaches for metadynamics reweighting, such as mean-force-integration schemes, have also been proposed.<sup>57</sup>

### Secondary-Structure Analysis, Solvent Exposure, Global Compaction, and Hydrogen Bonding

Per-residue secondary-structure assignments were obtained using the DSSP algorithm,<sup>54</sup> which classifies each residue at every frame into one of eight local conformations:  $3_{10}$  helix,  $\alpha$ -helix,  $\pi$ -helix, polyproline II, extended  $\beta$ -strand, isolated  $\beta$ -bridge, turn, and bend, with all remaining states reported as coil. For each frame, each DSSP label was encoded as a binary indicator and reweighted frame-by-frame as described in [Reweighting Procedure](#) Section, yielding unbiased per-residue probabilities  $p_i(s)$  for each DSSP state  $s$  at residue  $i$ .

Intra-peptide contact maps were computed using the MDAnalysis Python library.<sup>58,59</sup> Pairwise residue distances were calculated considering only  $C_\alpha$  atoms, and a cutoff of 0.8 nm was applied to define residue-residue contacts. This value is consistent with contact map analyses from previous A $\beta$  peptide studies, in which thresholds of 0.6–0.9 nm are used for  $C_\alpha$ -based definitions.<sup>41,60</sup> The robustness of the contact patterns was verified by repeating the analysis with  $C_\alpha$ – $C_\alpha$  cutoffs of 0.7 and 0.9 nm, which preserved the same qualitative isoform-specific trends. Contact probabilities were subsequently reweighted to obtain unbiased estimates.

Solvent-accessible surface area (SASA) was computed using `gmx sasa`,<sup>47</sup> which implements the double-cubic lattice method.<sup>61</sup> Total SASA was decomposed into polar (P-SASA) and apolar (A-SASA) contributions following the Roseman hydrophobicity scale:<sup>62</sup> polar residues contribute to P-SASA, apolar residues to A-SASA, and glycine is excluded from both categories. The C-terminal residue, modeled as deprotonated in both A $\beta$ 40 and A $\beta$ 42, was treated as polar regardless of its Roseman classification. All SASA-based quantities were reweighted to recover unbiased ensemble distributions.



**Figure 1.** *Aβ40* and *Aβ42* free energy surfaces. FES for (a) *Aβ40* and (b) *Aβ42* in the  $(n_\alpha, n_\beta)$  plane. The variables  $n_\alpha$  and  $n_\beta$  approximate the number of residues in  $\alpha$ -helical and  $\beta$ -like conformations, respectively, and highlight the different balance between helical and  $\beta$ -rich states in the two isoforms. Below, one-dimensional free energy profiles for (c) *Aβ40* and (d) *Aβ42*, projected onto  $n_\alpha$  and  $n_\beta$ . Solid lines show the reweighted free energy profiles, with shaded regions indicating the standard error of the mean across the three independent replicas.

Global compaction was quantified via the radius of gyration  $R_g$ , computed with `gmx polystat`<sup>47</sup> and subsequently reweighted.

Hydrogen bonds were evaluated using `gmx hbond`<sup>47</sup> with default angle and distance parameters, distinguishing intrapeptide ( $n_{\text{HB-IP}}$ ) and peptide–solvent hydrogen bonds ( $n_{\text{HB-PS}}$ ). Both time series were processed through the same reweighting procedure described above to obtain unbiased distributions.

### Conformational Clustering

In STEP 1 of the clustering procedure (Figure 5), we used  $n_{\text{max}} = 10$ ,  $s_{\text{min}} = 0.3$ , and  $N_{\text{min}} = 300$ .  $n_{\text{max}}$  was set to 10 as for  $n > 10$  we did not obtain acceptable clusters. The distance cutoff for local density calculation in DPC was initially set at 1.3 in the dimensional space of the four normalized descriptors and then reduced by 10% at each iteration. The iteration on the density cutoff stops when either the remaining frames are less than  $N_{\text{min}}$  or no acceptable cluster had been generated in the last 5 iterations.

In STEP 2, clusters produced by different subsets and density scales were compared using the Wasserstein distances.<sup>63,64</sup> In particular, as distances had to be calculated in the space of the four descriptors, we considered the modulus of the array composed of the four Wasserstein distances. The resulting distance matrix was subjected to agglomerative hierarchical clustering with average linkage.<sup>65</sup> The dendrogram was cut at the elbow of the merge-distance curve,<sup>66</sup> yielding a unified set of *global clusters* consistent across all subsets.

**Cluster Relevance.** As MetaD does not preserve unbiased state populations, raw cluster occupancies cannot be interpreted directly. To recover equilibrium populations, each frame was assigned a statistical weight  $w_k$  computed according to the reweighting formalism; see *Reweighting Procedure* Section. The relevance of a global cluster  $C_i$  was defined as

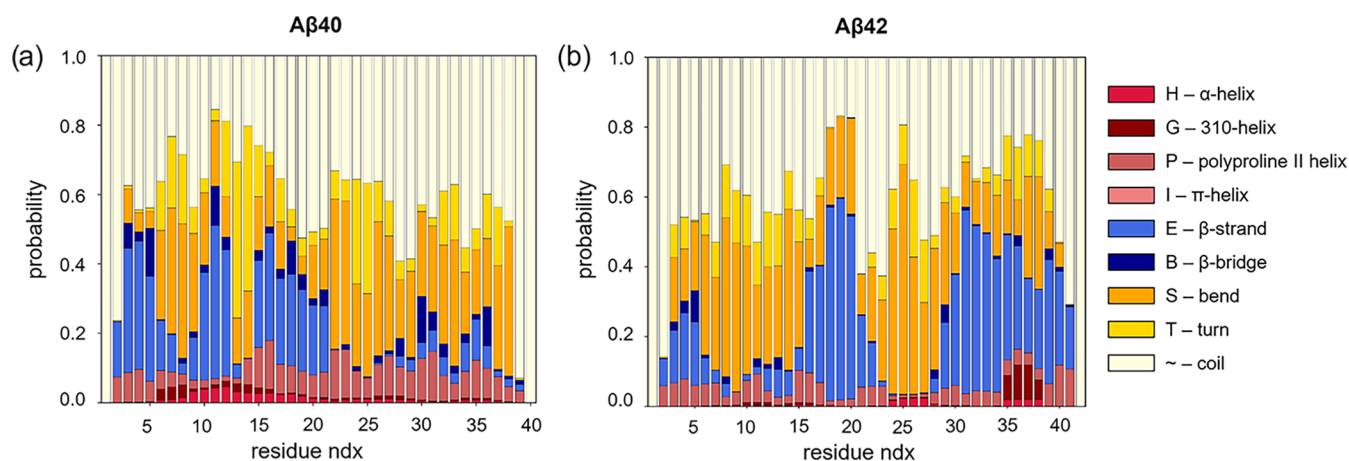
$$R_i = \frac{\sum_{k \in C_i} w_k}{\sum_k w_k}$$

which represents the fraction of the unbiased equilibrium probability density associated with cluster  $C_i$ . Accordingly, a cluster with high  $R_i$  corresponds to a structurally coherent and frequently sampled basin in the unbiased ensemble, whereas clusters with low  $R_i$  reflect rare or transient conformations.

This definition is also conceptually related to recent strategies in which conformer relevance or free energy differences are estimated by combining density-based conformational clustering with reweighting or multidimensional density-estimation procedures.<sup>67–69</sup> Overall, this framework integrates density-based detection, hierarchical merging, and replica-aware reweighting, yielding a statistically robust and physically interpretable partition of the *Aβ40* and *Aβ42* conformational landscapes.

## RESULTS

We characterized and compared the monomeric conformational ensembles of *Aβ40* and *Aβ42* using all-atom simulations with the CHARMM36m force field<sup>21</sup> in explicit water under near-physiological conditions (0.15 M NaCl, 310 K). CHARMM36m has been shown to reproduce conformational ensembles of intrinsically disordered peptides and proteins,<sup>21,39,46</sup> and CHARMM36m-based enhanced-sampling simulations have been successfully applied to *Aβ42* monomers in aqueous solution under physiological ionic strength.<sup>70</sup> As a first step, we reconstructed the free energy surfaces (FESs) of the two isoforms. We then characterized the resulting conformational ensembles in terms of residue-resolved secondary-structure propensities, global compaction, and



**Figure 2.** Per-residue secondary structure probabilities for A $\beta$ 40 (a) and A $\beta$ 42 (b). Each vertical bar corresponds to a residue, and the colored stacked segments represent the full probability distribution  $\{p_i(s)\}$  across eight DSSP categories. The total height of each bar is one, providing a direct readout of the relative sampling of each secondary structure type in the unbiased ensemble.

hydrophobic exposure, as well as the dominant metastable families identified through a tailored clustering procedure. See [Methods](#) Section and the [Supporting Information \(SI\)](#) for simulation details.

### Global Free Energy Surfaces

To obtain a global picture of the conformational preferences of the two monomers in water, we reconstructed the free energy surfaces (FESs) of A $\beta$ 40 and A $\beta$ 42 using well-tempered metadynamics. We used two collective variables (CVs) describing secondary structure content: the  $\alpha$ - and  $\beta$ -content, implemented in PLUMED as ALPHARMSD and ANTI-BETARMSD<sup>52</sup> and described in detail in the [Methods](#) Section. Convergence is a well-known critical aspect in the study of IDPs, whose conformational landscapes are broad and only weakly funneled. We monitored the stability of the sampling achieved in these simulations by monitoring the time evolution of the free energy difference between relevant FES regions (see [SI](#) for more details). The results indicate that the sampling is sufficiently stable to support the comparative analysis presented below for A $\beta$ 40 and A $\beta$ 42.

In order to present the results in a more easily interpretable collective variable space, the original CV values were rescaled to yield approximate residue counts for  $\alpha$ -helical ( $n_\alpha$ ) and  $\beta$ -sheet ( $n_\beta$ ) structures. The FESs were therefore represented in the  $(n_\alpha, n_\beta)$  plane for each isoform. The resulting two-dimensional FESs and their corresponding one-dimensional projections are shown in [Figure 1](#).

The FESs of both peptides reflect their intrinsically disordered nature, displaying the characteristic inverted funnel shape in which highly disordered states occupy the lowest free energy region, while more ordered conformations enriched in secondary structure elements correspond to higher free energy values.<sup>36,71</sup> Both isoforms span broad, multimodal conformational spaces. The most populated regions correspond to poorly structured, coil-like states, with additional basins associated with significant  $\beta$ -sheet content.

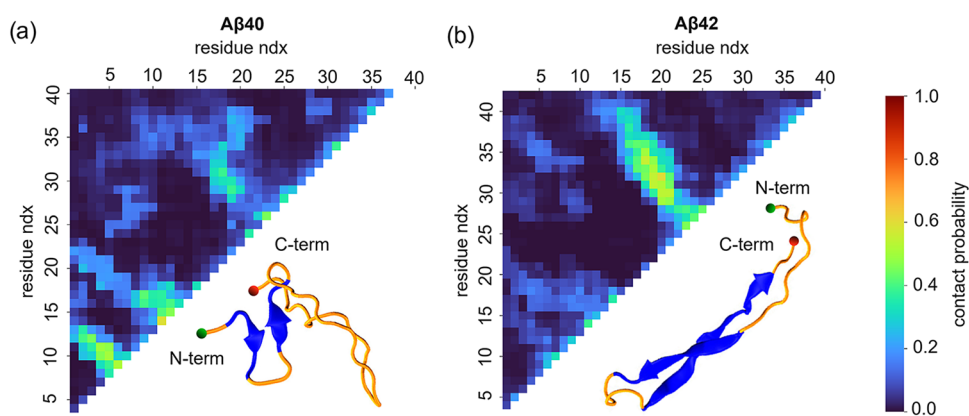
Across both peptides,  $\beta$ -like conformations are generally favored over  $\alpha$ -helical ones. A $\beta$ 42 samples  $\beta$ -rich regions more extensively than A $\beta$ 40, whereas A $\beta$ 40 explores the landscape more broadly, visiting low and medium  $\alpha$ -content regions more frequently, although these remain at high free energy values.

### Local Secondary Structure Patterns and Sequence-Dependent Differences

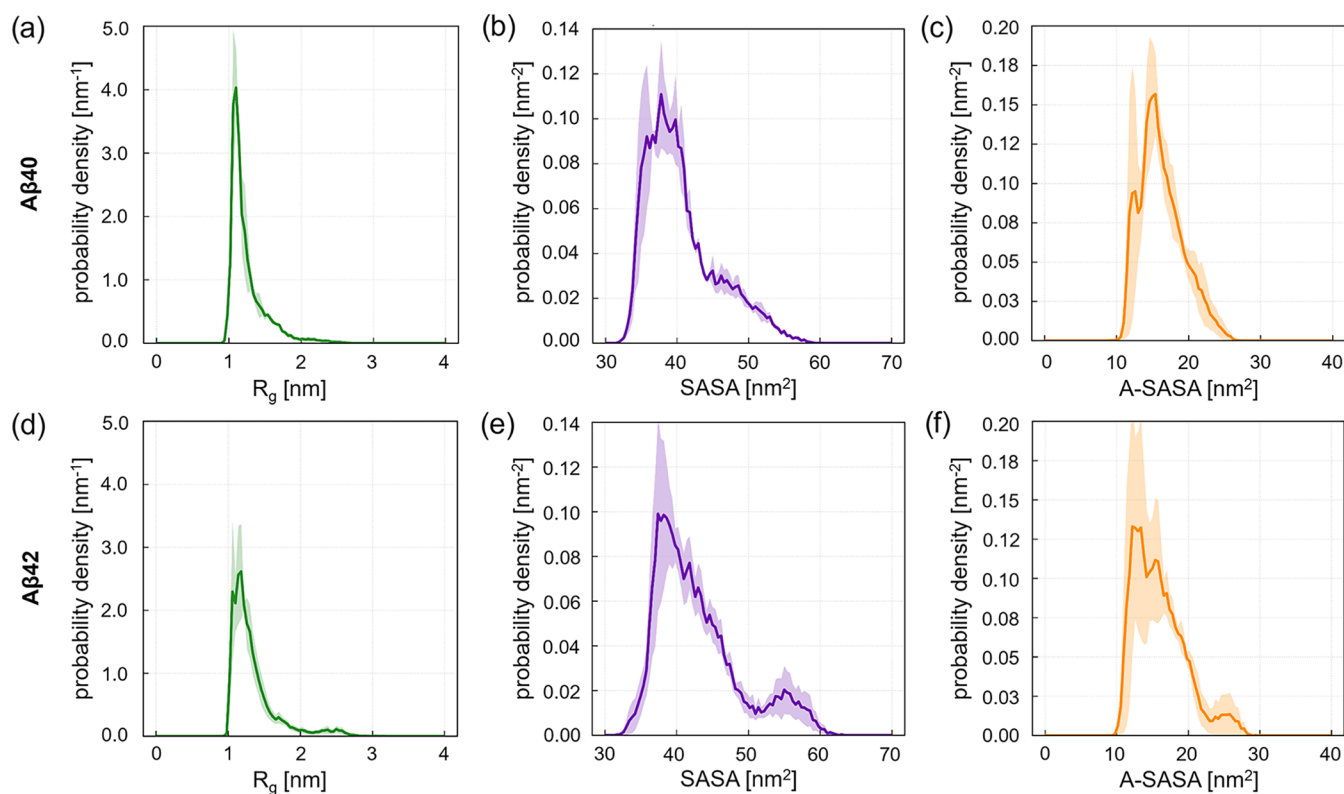
Per-residue secondary structure probabilities ([Figure 2](#)) were obtained by reweighting (see [Methods](#) Section) and analyzing each trajectory with DSSP (version 3.1.4).<sup>54</sup> For every frame and every residue, DSSP assigns one among eight local conformations—i.e.,  $3_{10}$  helix,  $\alpha$ -helix,  $\pi$ -helix, polyproline II helix, extended  $\beta$ -strand, isolated  $\beta$ -bridge, turn, and bend—with remaining states reported as coil. Each assignment was encoded as a binary indicator and reweighted frame-by-frame, yielding unbiased probabilities  $p_i(s)$  for all DSSP states  $s$  at each residue  $i$ . In [Figure 2](#), the probability distributions  $\{p_i(s)\}$  across the eight DSSP categories are shown for each residue. Results for the individual replicas are reported in the [Supporting Information](#), confirming the robustness and reproducibility of the observed secondary-structure patterns. For both peptides, as expected, coil-like states dominate, but pronounced sequence-dependent differences emerge in the pattern and localization of  $\beta$ -strand formation.

A $\beta$ 40 displays three well-separated  $\beta$  peaks in the first half of the sequence (approximately residues 2–7, 9–13, and 15–21), each sharp and localized, with intervening residues reverting almost entirely to coil or turn states. Beyond residue  $\sim 21$ , in the C-terminal region,  $\beta$  structures are only sparsely observed.

A $\beta$ 42, by contrast, shows a distinctly different organization. In the N-terminal half, the first  $\beta$  peak (residues 2–7) is present but weaker than in A $\beta$ 40, and the second peak (9–13) is nearly absent. The third peak (15–21) is instead present and relevant. The most striking feature is a broad, continuous, and strongly stabilized  $\beta$ -rich region extending approximately from residue 29 to the C-terminus. This persistent C-terminal  $\beta$ -segment has no counterpart in A $\beta$ 40 and is consistent with independent simulations and NMR-refined ensembles highlighting enhanced  $\beta$ -structure in the 30–36 and 39–41 regions of A $\beta$ 42.<sup>7,34,72</sup> Notably, the third peak (residues 15–21) corresponds to the central hydrophobic core (CHC), a region extensively identified as a key self-recognition and nucleation motif in A $\beta$  aggregation and identified as one of the  $\beta$ -strands constituting the fibril core of full-length A $\beta$ .<sup>73,74</sup> The significant  $\beta$ -peak observed in the 15–21 region for both isoforms indicates that the CHC already samples  $\beta$ -rich conformations at the monomeric level, potentially lowering the free energy



**Figure 3.** Intra-peptide contact maps and representative structures sampled for (a)  $A\beta_{40}$  and (b)  $A\beta_{42}$ . Each square represents the contact probability between pairs of residues. The main diagonal is masked to exclude trivial contacts between sequentially adjacent residues. Besides the contact maps, an example representative snapshot of the corresponding peptide is shown. The N-terminal and C-terminal ends of both peptides are indicated by green and red beads, respectively. Blue arrows highlight antiparallel  $\beta$ -sheets.



**Figure 4.** Conformations and exposure to solvents of (a–c)  $A\beta_{40}$  and (d–f)  $A\beta_{42}$ . (a, d) Probability distributions of the radius of gyration ( $R_g$ ). (b, e) Total solvent-accessible surface area (SASA). (c, f) Apolar SASA (A-SASA). Shaded regions represent the standard error of the mean across independent replicas.

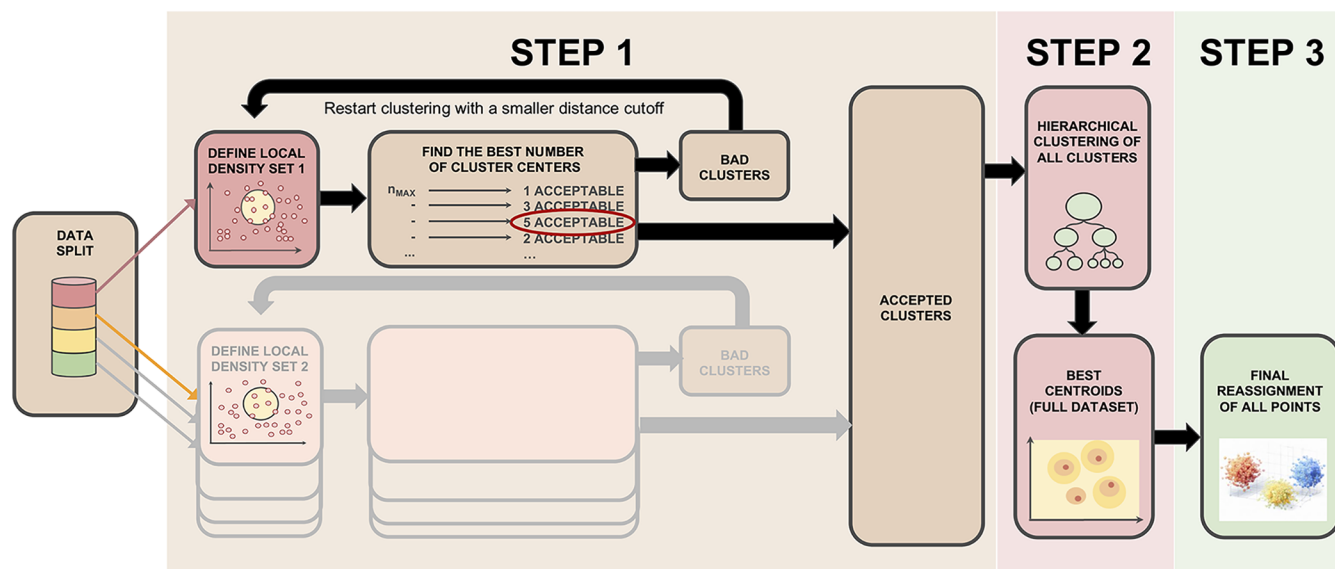
barrier for nucleation and supporting its role as an intrinsic aggregation hotspot.

Additional DSSP features contextualize these differences. Turns and bends are abundant in both isoforms and preferentially cluster near the boundaries of the  $\beta$  regions. Helical conformations are overall rare but sampled slightly more frequently in  $A\beta_{40}$ , indicating a more heterogeneous and structurally flexible ensemble.

We conducted a complementary analysis by calculating intra-peptide contact maps (Figure 3) by reweighting and analyzing each trajectory with the MDAnalysis Python library.<sup>58</sup> Pairwise residue distances were calculated consider-

ing only  $C_\alpha$  atoms, and a cutoff of 0.8 nm was applied to define residue–residue contacts. The contact maps are consistent with the trends observed in the secondary-structure analysis (Figure 2) and offer a detailed, quantitative characterization of the peptides' dynamic organization. Replica-resolved contact maps are provided in the Supporting Information and show consistent trends across independent simulations.

For  $A\beta_{40}$  (Figure 3a), the highest contact probabilities are localized within the first half of the peptide sequence (residues 2–16), consistent with the local folding and stabilization of short and transient compact structures. There is another shallow contact region involving residues in the central part of



**Figure 5.** Schematic overview of the conformational clustering workflow. The procedure, described in detail in the main text, combines iterative density-based clustering and consensus consolidation to identify robust conformational families in heterogeneous ensembles.

the peptide; in contrast, the C-terminal region (29–40) shows no significant propensity to establish persistent contacts with other segments of the peptide.

Conversely, for  $A\beta_{42}$  (Figure 3b), the C-terminal region exhibits the highest contact probabilities, particularly through interactions with the central hydrophobic core (residues 17–21). Our results are in agreement with previous experimental and computational studies, which have shown that the CHC, owing to its strong hydrophobic character and intrinsic  $\beta$ -strand propensity, promotes intermolecular cross-interactions already during the earliest stage of self-assembly and constitutes one of the principal structural elements of the amyloid fibril core. More specifically, CHC can participate in  $\beta$ -hairpin conformations involving the C-terminal region, particularly in  $A\beta_{42}$ , where interactions between residues in the CHC and the hydrophobic C-terminus stabilize aggregation-prone conformations.<sup>60,75</sup> Such  $\beta$ -hairpin motifs are considered important intermediates in early oligomerization, as they maintain extended  $\beta$ -strand-like conformations and expose hydrophobic side chains that facilitate intermolecular association and cross- $\beta$  assembly. The relative absence of analogous persistent contacts in  $A\beta_{40}$  supports the view that the weaker coupling between the CHC and the C-terminal region contributes to its lower aggregation propensity and greater conformational heterogeneity.<sup>72</sup> These observations align with previous studies<sup>76,77</sup> that relate the enhanced  $\beta$ -sheet propensity, the presence of two additional residues, and the higher structural rigidity of the  $A\beta_{42}$  C-terminal, and suggest a higher aggregation propensity compared to the more flexible  $A\beta_{40}$ .

### Compaction and Solvent Exposure

$A\beta_{40}$  and  $A\beta_{42}$  exhibit similar  $R_g$  distributions, both dominated by a main peak in the 1.0–1.5 nm range (Figure 4a,d), in line with previous all-atom simulations of  $A\beta$  monomers in explicit solvent.<sup>41,78</sup> This main peak corresponds to compact conformations representing the dominant population of both isoforms. However,  $A\beta_{42}$  displays a more evident tail extending up to  $\sim 2.5$ –3.0 nm, whereas  $A\beta_{40}$  decays more rapidly beyond 2 nm.

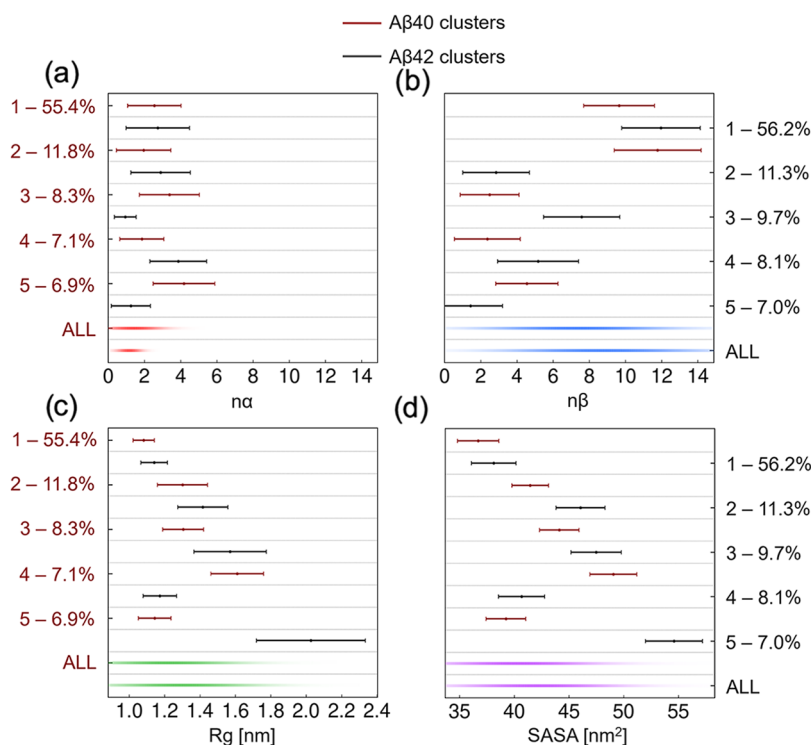
Consistent with these observations, the SASA distributions of both peptides show a primary peak around  $\approx 40$  nm<sup>2</sup> (Figure 4b,e). For  $A\beta_{40}$ , this peak is followed by a gentle, featureless shoulder at larger SASA values, reflecting a continuum of moderately more exposed conformations. In contrast,  $A\beta_{42}$  exhibits a second, distinct peak in the 50–60 nm<sup>2</sup> range, marking a well-defined population of highly solvent-exposed structures.

In addition to the total SASA, apolar SASA (A-SASA) was computed from the exposure of hydrophobic residues according to the Roseman scale<sup>62</sup> (See Methods Section for further details). A similar trend emerges for A-SASA (Figure 4c,f). Both peptides show a main peak between 15 and 20 nm<sup>2</sup>, corresponding to compact states with partially buried hydrophobic residues. Only  $A\beta_{42}$  displays an additional shoulder or minor peak between 20 and 30 nm<sup>2</sup>, indicating access to conformations with substantially larger hydrophobic exposure.

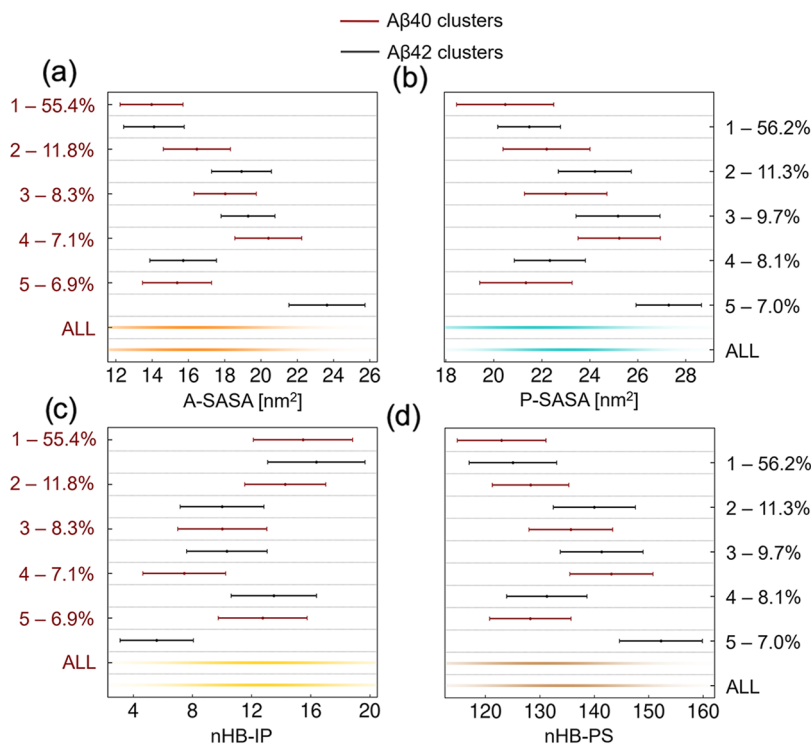
### Conformational Clustering

To resolve the dominant conformational families underlying the free energy surfaces of  $A\beta_{40}$  and  $A\beta_{42}$ , we developed an *ad hoc* clustering framework. Standard clustering approaches often struggle when applied to intrinsically disordered peptides, whose free energy landscapes are characterized by broad and partially overlapping basins.<sup>69,79</sup> When projected onto a low-dimensional space defined by a small set of CVs, these basins often appear as clusters of configurations with highly heterogeneous local densities.

Clustering was performed in the space of four structural descriptors, namely  $n_\alpha$ ,  $n_\beta$ , the radius of gyration  $R_g$ , and total SASA. We emphasize that our goal is not to provide a fully exhaustive structural classification, which is difficult to define unambiguously for intrinsically disordered peptides, but rather to identify and compare the main conformational basins sampled by the system. These variables were selected based on their pronounced multimodal behavior in the reweighted ensemble, indicating their ability to distinguish between different regions of the conformational space. In this context,  $n_\alpha$  and  $n_\beta$  describe the secondary structure content, while  $R_g$  and SASA provide complementary information on the overall



**Figure 6.** Characterization of the dominant structural families in terms of the structural descriptors used in the consensus clustering analysis. Panels (a–d) report, respectively, the structural descriptors  $n_{\alpha}$ ,  $n_{\beta}$ ,  $R_g$ , and SASA, which define the clustering space. In each panel, the five most relevant clusters of A $\beta$ 40 and A $\beta$ 42 are shown, ordered by decreasing relevance, with the corresponding reweighted population reported in parentheses. For each cluster and descriptor, the weighted mean value is reported together with a horizontal bar representing the width of the reweighted distribution. For reference, background distributions corresponding to the full reweighted ensembles of A $\beta$ 40 and A $\beta$ 42 are shown, with color intensity proportional to the probability density.



**Figure 7.** Characterization of the dominant structural families in terms of structural descriptors excluded from the consensus clustering analysis. Panels (a–d) report, respectively, A-SASA, P-SASA,  $n_{\text{HB-IP}}$ , and  $n_{\text{HB-PS}}$ , computed *a posteriori* and not used for clustering. All other conventions are as in Figure 6.

size, compactness, and solvent exposure of the peptide. These features are also relevant to aggregation propensity, as  $\beta$ -rich content, compaction, and solvent exposure are commonly associated with early stages of amyloid formation.

The clustering strategy adopted here was therefore specifically designed to address these challenges by combining density-based detection across multiple scales with a consensus-based reconciliation of independently clustered trajectory subsets.<sup>80</sup> A schematic overview of the workflow is shown in Figure 5, while the full workflow is explained below.

**Consensus Clustering Workflow.** To improve statistical robustness and control memory usage, trajectories were merged and partitioned into random subsets of about  $3 \times 10^4$  frames. We verified that the distributions of structural descriptors were essentially indistinguishable across the different subsets.

The first step of our clustering approach (STEP 1 in Figure 5) applies a parallel workflow to each subset. The workflow is based on the standard DPC algorithm,<sup>81</sup> which identifies cluster centers as points of simultaneously high density and large distance from any denser point. Similar density-peak-based strategies have recently been applied to the unsupervised identification of molecular conformers and conformational states.<sup>69,79</sup> DPC is applied recursively, each time reducing the number of requested clusters from  $n_{\max}$  to 2. As a result of this scan, the optimal  $n$  is identified as that generating the highest number of acceptable clusters (silhouette score  $\geq s_{\min}$  and size  $N \geq N_{\min}$  frames). Acceptable clusters are then removed from the subset sample and saved. The discarded frames are resubmitted to DPC with a reduced distance cutoff (see Methods Section for details). The procedure is iterated until no acceptable clusters are identified. At the end of STEP 1, it is possible that some frames remain unassigned.

In STEP 2 (Figure 5), all acceptable clusters identified at STEP 1 from all the trajectory subsets are subject to a hierarchical clustering procedure, based on the Wasserstein (Earth Mover's) distance and agglomerative hierarchical clustering with average linkage,<sup>65</sup> as better detailed in the Methods Section. The output of STEP 2 is the final list of centroids.

In STEP 3, all the original trajectory frames, including those that were not assigned in STEP 1, are reassigned to their closest centroid.

The relevance of each conformational family was computed *a posteriori* by incorporating the reweighting factors, thus providing an unbiased estimate of its equilibrium population. The specific parameters used in the clustering pipeline, together with additional technical details, are provided in the Methods Section, while the SI reports the full structural characterization for all clusters.

**Outcomes.** Figures 6 and 7 summarize the structural characterization of the dominant conformational families identified by the consensus clustering for A $\beta$ 40 and A $\beta$ 42. Figure 6 compares clusters using the descriptors defining the clustering space, while Figure 7 provides an additional *a posteriori* characterization based on complementary observables—i.e., A-SASA, P-SASA,  $n_{\text{HB-IP}}$ , and  $n_{\text{HB-PS}}$ . In both figures, clusters are reported in decreasing order of relevance. Throughout the manuscript, cluster indices (e.g., Cluster 1, Cluster 2) refer to the order of relevance for the corresponding peptide, as defined in Figures 6 and 7.

The dominant conformational families identified by the consensus clustering reveal clear structural correspondences as

well as systematic differences between A $\beta$ 40 and A $\beta$ 42. A detailed comparison of the most relevant clusters highlights how the two isoforms populate similar basins with distinct degrees of compactness, secondary structure content, and solvent exposure:

**Dominant Compact  $\beta$ -Rich Basin.** Cluster 1 in both peptides corresponds to compact,  $\beta$ -rich conformations that dominate the equilibrium ensemble. Notably, A $\beta$ 42 exhibits a slightly higher degree of internal organization than A $\beta$ 40, as reflected by its larger number of intrapeptide hydrogen bonds.

**Extended  $\beta$ -Rich Conformations.** A clear correspondence emerges between Cluster 2 of A $\beta$ 40 and Cluster 3 of A $\beta$ 42. These states represent more extended and slightly less  $\beta$ -rich conformations with respect to Cluster 1. In this regime, A $\beta$ 40 displays a higher  $\beta$ -structure content, whereas A $\beta$ 42 adopts more expanded conformations, characterized by increased  $R_g$ , SASA, A-SASA, and P-SASA. Consistently, A $\beta$ 40 retains a larger number of intrapeptide hydrogen bonds, while A $\beta$ 42 interacts more extensively with the solvent.

**Highly Disordered Compact States.** A second pair of corresponding states is captured by Cluster 3 of A $\beta$ 40 and Cluster 2 of A $\beta$ 42. Both clusters are highly disordered while remaining relatively compact; nevertheless, A $\beta$ 42 adopts more open conformations and displays higher SASA values.

**Mixed and Weakly  $\alpha$ -Structured Compact Ensembles.** Cluster 5 of A $\beta$ 40 and Cluster 4 of A $\beta$ 42 represent the only basins with detectable, albeit low,  $\alpha$ -helical content. These clusters exhibit comparable  $\beta$ -levels and correspond to mixed, compact, and internally organized conformational ensembles.

**Expanded and Solvent-Exposed Disordered States.** Finally, Cluster 4 of A $\beta$ 40 and Cluster 5 of A $\beta$ 42 correspond to more disordered and expanded states with increased solvent exposure. The A $\beta$ 42 basin is significantly more extended and solvent-exposed, in agreement with the bimodal behavior observed in the SASA and A-SASA distributions (Figure 4).

## DISCUSSION AND CONCLUSION

By combining molecular dynamics simulations with enhanced sampling techniques, this work aimed to determine whether monomer-level conformational biases could already provide a structural basis for the distinct aggregation propensities observed for A $\beta$ 40 and A $\beta$ 42. While many aspects of the conformational ensembles of these two peptides have been examined previously, here we provide a direct comparison obtained under identical simulation and analysis conditions (same sampling strategy, force field, and postprocessing pipeline).

The reconstructed free energy surfaces (Figure 1) show that both peptides predominantly populate disordered, coil-like ensembles, in agreement with previous computational studies.<sup>36,71</sup> Nevertheless, clear differences emerge in their secondary structure balance. A $\beta$ 42 exhibits a higher probability of adopting  $\beta$ -like conformations relative to  $\alpha$ -helical ones, exceeding the corresponding tendency observed in A $\beta$ 40. Conversely, A $\beta$ 40 samples a broader range of conformations, including more frequent excursions into transient  $\alpha$ -helical regions. These global trends indicate a subtle but systematic shift in the balance between disorder and transient structural order that is already encoded at the monomeric level.

The residue-resolved DSSP analysis (Figure 2) refines this picture by identifying the sequence regions responsible for these global differences. In particular, the extended C-terminal  $\beta$  domain observed in A $\beta$ 42 closely mirrors experimentally

established distinctions between the fibril cores of the two peptides. Solid-state NMR and cryo-EM studies consistently identify residues 30–42, together with a central segment around residues 10–20, as major contributors to the cross- $\beta$  core in  $A\beta$ 42 fibrils, whereas  $A\beta$ 40 incorporates C-terminal residues only weakly, or not at all, under comparable conditions.<sup>43,44,83–85</sup> Recent meta-analyses by Errico et al.<sup>18</sup> further show that, across dozens of *in vitro* and brain-derived  $A\beta$  fibril polymorphs, dominant cross- $\beta$  cores almost invariably involve overlapping C-terminal segments, despite their high structural diversity. Although the present simulations focus on monomeric species, the presence of such motifs suggests that the  $A\beta$ 42 monomeric ensemble samples conformations that are structurally consistent with aggregation-prone features identified experimentally. In contrast, the more fragmented and heterogeneous  $\beta$ -strand distribution observed in  $A\beta$ 40 is associated with a conformational landscape that appears less prone to forming the contiguous  $\beta$  segments characteristic of mature fibrils. Consistently, the intrapeptide contact maps reveal an isoform-specific reorganization of tertiary contacts (Figure 3):  $A\beta$ 42 displays enhanced contact probability between the C-terminal segment and the central hydrophobic core, whereas  $A\beta$ 40 shows contacts mainly confined to the N-terminal/central region with little persistent involvement of its C-terminus.

Global shape and solvent exposure provide a complementary perspective on the monomeric ensembles (Figure 4). In our simulations,  $A\beta$ 42 displays a slightly broader distribution of the radius of gyration compared to  $A\beta$ 40, together with a distinct secondary peak in the SASA and A-SASA distributions associated with extended, highly solvent-exposed conformations. By contrast, the corresponding distributions for  $A\beta$ 40 decay more rapidly toward larger radii and higher solvent exposure. These  $A\beta$ 42-specific shoulders indicate that, while both isoforms share similarly most-probable compact states,  $A\beta$ 42 more readily samples conformations with larger surface area and greater hydrophobic exposure.

Previous experimental and computational studies have reported partially divergent views on the conformational differences between  $A\beta$ 40 and  $A\beta$ 42. Single-molecule FRET measurements combined with molecular dynamics simulations by Meng et al.<sup>15</sup> suggested that both isoforms populate highly disordered ensembles with nearly indistinguishable ensemble-averaged radii of gyration and end-to-end distances. In contrast, extensive all-atom simulations by Song et al.<sup>38</sup> reported that, while both peptides sample collapsed and extended states,  $A\beta$ 42 exhibits an enhanced propensity to form compact, hydrophobically collapsed conformations with reduced hydrophobic SASA compared to  $A\beta$ 40. Taken together, these apparently conflicting observations highlight that ensemble-averaged descriptors such as  $R_g$  and SASA are highly sensitive to force-field choices, modeling resolution, and analysis protocols, and may therefore mask substantial heterogeneity within the conformational ensemble. Rather than indicating an intrinsic ambiguity, this sensitivity underscores the need for analysis strategies capable of resolving recurrent conformational families beyond simple averages.

In this context, the clustering framework introduced in this study enables a more robust and physically interpretable comparison of the conformational landscapes of  $A\beta$ 40 and  $A\beta$ 42 by explicitly identifying and contrasting their dominant structural basins. The cluster-level analysis (Figures 6 and 7) confirms the trends inferred from global observables, showing

that  $A\beta$ 42 displays both a slightly higher degree of internal organization in  $\beta$ -rich states and a greater propensity to populate extended, solvent-exposed disordered basins compared to  $A\beta$ 40.

Taken together, our results demonstrate that the addition of just two residues in  $A\beta$ 42 is sufficient to bias the monomeric conformational ensemble toward a shifted balance between disorder and transient order. This shift is characterized by an increased propensity for  $\beta$ -like structures, a stronger involvement of the C-terminal region, and more frequent sampling of extended and solvent-exposed conformations. These monomeric preferences provide a possible structural rationale for the early stages of the aggregation cascade. The  $\beta$ -rich basins identified in our free energy landscapes can be related to the “assembly-competent” conformational states ( $N^*$ states) proposed in previous monomer studies as precursors for oligomer formation.<sup>34</sup> In particular, the persistent  $\beta$ -rich segment observed in the  $A\beta$ 42 C-terminal region (approximately residues 29–42), together with its enhanced interaction propensity with the central hydrophobic core (CHC), is consistent with structural motifs reported for aggregation-prone dimers.<sup>86</sup> Previous simulations have shown that dimerization promotes intermolecular  $\beta$ -sheet formation involving these same sequence regions, accompanied by a substantial increase in  $\beta$ -sheet content relative to the monomeric state.<sup>10,86,87</sup> The bimodal SASA and  $R_g$  distributions observed specifically for  $A\beta$ 42 may also have important implications for higher-order oligomerization. The increased population of extended and solvent-exposed conformations likely facilitates intermolecular encounters during the early stages of self-assembly. In this respect, our results are consistent with previous reports showing that  $A\beta$ 42 dimers and tetramers exhibit larger hydrophobic solvent-accessible surface areas (hSASA) than the corresponding  $A\beta$ 40 assemblies.<sup>34,88,89</sup> Such enhanced hydrophobic exposure has been proposed to favor the recruitment of additional monomers, promote the formation of compact oligomeric species, and contribute to the increased cytotoxicity associated with  $A\beta$ 42 oligomers.<sup>89</sup>

More broadly, our findings suggest that the divergent aggregation behaviors of  $A\beta$ 40 and  $A\beta$ 42 may originate from small but systematic redistributions of conformational populations already present at the monomeric level. Such subtle biases can be amplified during oligomerization and fibril formation, highlighting how minimal sequence variations in intrinsically disordered peptides can produce disproportionately large effects on collective behavior.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.6c01000>.

WT-MetaD convergence assessment; validation of PLUMED-based secondary-structure estimators against DSSP and STRIDE; complete cluster-resolved characterization of  $A\beta$ 40 and  $A\beta$ 42 conformational ensembles; per-residue secondary structure probabilities, per-residue intrapeptide contact maps (PDF)

## AUTHOR INFORMATION

## Corresponding Authors

Valeria Conti Nibali – Department of Mathematical and Computational Sciences, Physical Sciences and Earth Sciences, University of Messina, Messina 98166, Italy; Email: [valeria.continibali@unime.it](mailto:valeria.continibali@unime.it)

Davide Bochicchio – Department of Physics, University of Genoa, Genoa 16146, Italy; [orcid.org/0000-0002-3682-9086](https://orcid.org/0000-0002-3682-9086); Email: [davide.bochicchio@unige.it](mailto:davide.bochicchio@unige.it)

## Authors

Irene Cadenelli – Department of Physics, University of Genoa, Genoa 16146, Italy

Andrea Ciccolo – Department of Mathematical and Computational Sciences, Physical Sciences and Earth Sciences, University of Messina, Messina 98166, Italy

Andrea Tagliabue – Department of Physics, University of Genoa, Genoa 16146, Italy; [orcid.org/0000-0001-9520-0627](https://orcid.org/0000-0001-9520-0627)

Giulia Rossi – Department of Physics, University of Genoa, Genoa 16146, Italy; [orcid.org/0000-0001-6916-2049](https://orcid.org/0000-0001-6916-2049)

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpccb.6c01000>

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was financially supported by the Italian Ministry of University and Research (MUR) through the PRIN 2022 program, project NeuRaftAmyl, grant/project code 20222YMARS.

## REFERENCES

- (1) Panza, F.; Lozupone, M.; Logroscino, G.; Imbimbo, B. P. *Nat. Rev. Neurol.* **2019**, *15*, 73–88.
- (2) Korczyn, A. D.; Grinberg, L. T. *Nat. Rev. Neurol.* **2024**, *20*, 245–251.
- (3) Sehar, U.; Rawat, P.; Reddy, A. P.; Kopel, J.; Reddy, P. H. *Int. J. Mol. Sci.* **2022**, *23*, No. 12924.
- (4) Villegas, S.; Roda, A.; Serra-Mir, G.; Montoliu-Gaya, L.; Tiessler, L. *Neural Regen. Res.* **2022**, *17*, 1666–1674.
- (5) Heneka, M. T.; van der Flier, W. M.; Jessen, F.; Hoozemans, J.; Thal, D. R.; Boche, D.; Brosseron, F.; Teunissen, C.; Zetterberg, H.; Jacobs, A. H.; Edison, P.; Ramirez, A.; Cruchaga, C.; Lambert, J.-C.; Laza, A. R.; Sanchez-Mut, J. V.; Fischer, A.; Castro-Gomez, S.; Stein, T. D.; Kleineidam, L.; Wagner, M.; Neher, J. J.; Cunningham, C.; Singhrao, S. K.; Prinz, M.; Glass, C. K.; Schlachetzki, J. C. M.; Butovsky, O.; Kleemann, K.; De Jaeger, P. L.; Scheiblich, H.; Brown, G. C.; Landreth, G.; Moutinho, M.; Grutzendler, J.; Gomez-Nicola, D.; McManus, R. M.; Andreasson, K.; Ising, C.; Karabag, D.; Baker, D. J.; Liddelow, S. A.; Verkhratsky, A.; Tansey, M.; Monsonego, A.; Aigner, L.; Dorothée, G.; Nave, K.-A.; Simons, M.; Constantin, G.; Rosenzweig, N.; Pascual, A.; Petzold, G. C.; Kipnis, J.; Venegas, C.; Colonna, M.; Walter, J.; Tenner, A. J.; O'Banion, M. K.; Steinert, J. R.; Feinstein, D. L.; Sastre, M.; Bhaskar, K.; Hong, S.; Schafer, D. P.; Golde, T.; Ransohoff, R. M.; Morgan, D.; Breitner, J.; Mancuso, R.; Riechers, S.-P.; et al. *Nat. Rev. Immunol.* **2025**, *25*, 321–352.
- (6) Cohen, S. I. A.; Linse, S.; Luheshi, L. M.; Hellstrand, E.; White, D. A.; Rajah, L.; Otzen, D. E.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 9758–9763.
- (7) Meisl, G.; Yang, X.; Hellstrand, E.; Frohm, B.; Kirkegaard, J. B.; Cohen, S. I. A.; Dobson, C. M.; Linse, S.; Knowles, T. P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 9384–9389.
- (8) Qiu, T.; Liu, Q.; Chen, Y.; Zhao, Y.; Li, Y. J. *Pept. Sci.* **2015**, *21*, 522–529.
- (9) Nguyen, P. H.; Derreumaux, P. *Molecules* **2023**, *28*, No. 7080.
- (10) Nguyen, P. H.; Ramamoorthy, A.; Sahoo, B. R.; Zheng, J.; Faller, P.; Straub, J. E.; Dominguez, L.; Shea, J.-E.; Dokholyan, N. V.; De Simone, A.; Ma, B.; Nussinov, R.; Najafi, S.; Ngo, S. T.; Loquet, A.; Chiriccotto, M.; Ganguly, P.; McCarty, J.; Li, M. S.; Hall, C.; Wang, Y.; Miller, Y.; Melchionna, S.; Habenstein, B.; Timr, S.; Chen, J.; Hnath, B.; Strodel, B.; Kaye, R.; Lesné, S.; Wei, G.; Sterpone, F.; Doig, A. J.; Derreumaux, P.; et al. *Chem. Rev.* **2021**, *121*, 2545–2647.
- (11) Nguyen, H. L.; Krupa, P.; Hai, N. M.; Linh, H. Q.; Li, M. S. J. *Phys. Chem. B* **2019**, *123*, 7253–7269.
- (12) Thacker, D.; Barghouth, M.; Bless, M.; Zhang, E.; Linse, S. *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120*, No. e2220664120, DOI: [10.1073/pnas.2220664120](https://doi.org/10.1073/pnas.2220664120).
- (13) Bunce, S. J.; Wang, Y.; Stewart, K. L.; Ashcroft, A. E.; Radford, S. E.; Hall, C. K.; Wilson, A. J. *Sci. Adv.* **2019**, *5*, No. eaav8216, DOI: [10.1126/sciadv.aav8216](https://doi.org/10.1126/sciadv.aav8216).
- (14) Nguyen, P. H.; Sterpone, F.; Derreumaux, P. *Proteins: Struct., Funct., Bioinf.* **2025**, *93*, 1308–1313.
- (15) Meng, F.; Bellaiche, M. M.; Kim, J.-Y.; Zerze, G. H.; Best, R. B.; Chung, H. S. *Biophys. J.* **2018**, *114*, 870–884.
- (16) Rosenman, D. J.; Connors, C. R.; Chen, W.; Wang, C.; Garcia, A. E. J. *Mol. Biol.* **2013**, *425*, 3338–3359.
- (17) Zheng, W.; Tsai, M.-Y.; Wolynes, P. G. *J. Am. Chem. Soc.* **2017**, *139*, 16666–16676.
- (18) Errico, S.; Fani, G.; Ventura, S.; Schymkowitz, J.; Rousseau, F.; Trovato, A.; Vendruscolo, M.; Bemporad, F.; Chiti, F. *Biochem. J.* **2025**, *482*, 87–101.
- (19) Ni, C.; Shi, H.; Yu, H.; Chang, Y.; Chen, Y. *FASEB J.* **2011**, *25*, 1390–1401.
- (20) Teplow, D. B. *Amyloid, Prions, and Other Protein Aggregates, Part C, Methods in Enzymology*; Elsevier, 2006; Vol. 413, pp 20–33.
- (21) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. *Nat. Methods* **2017**, *14*, 71–73.
- (22) Robustelli, P.; Piana, S.; Shaw, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E4758–E4766, DOI: [10.1073/pnas.1800690115](https://doi.org/10.1073/pnas.1800690115).
- (23) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. *J. Phys. Chem. B* **2015**, *119*, 5113–5123.
- (24) Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 6817–6822.
- (25) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, No. 020603, DOI: [10.1103/PhysRevLett.100.020603](https://doi.org/10.1103/PhysRevLett.100.020603).
- (26) Tiwary, P.; Parrinello, M. *J. Phys. Chem. B* **2015**, *119*, 736–742.
- (27) Marques, S. M.; Kouba, P.; Legrand, A.; Sedlar, J.; Disson, L.; Planas-Iglesias, J.; Sanusi, Z.; Kunka, A.; Damborsky, J.; Pajdla, T.; Prokop, Z.; Mazurenko, S.; Sivic, J.; Bednar, D. *JACS Au* **2024**, *4*, 2228–2245.
- (28) Löhr, T.; Kohlhoff, K.; Heller, G. T.; Camilloni, C.; Vendruscolo, M. *Nat. Comput. Sci.* **2021**, *1*, 71–78.
- (29) Huang, Y.; Zhang, H.; Lin, Z.; Wei, Y.; Xi, W. *Methods* **2024**, *229*, 163–174.
- (30) Janson, G.; Valdes-Garcia, G.; Heo, L.; Feig, M. *Nat. Commun.* **2023**, *14*, No. 774, DOI: [10.1038/s41467-023-36443-x](https://doi.org/10.1038/s41467-023-36443-x).
- (31) Janson, G.; Feig, M. *PLOS Comput. Biology* **2024**, *20*, No. e1012144.
- (32) Ruzmetov, T.; Hung, T. I.; Jonnalagedda, S. P.; Chen, S.-h.; Fasihianifard, P.; Guo, Z.; Bhanu, B.; Chang, C.-e. A. *J. Chem. Inf. Model.* **2025**, *65*, 2487–2502.
- (33) Roche, J.; Shen, Y.; Lee, J. H.; Ying, J.; Bax, A. *Biochemistry* **2016**, *55*, 762–775.
- (34) Chakraborty, D.; Straub, J. E.; Thirumalai, D. *Sci. Adv.* **2023**, *9*, No. eadd6921, DOI: [10.1126/sciadv.add6921](https://doi.org/10.1126/sciadv.add6921).
- (35) Sgourakis, N. G.; Yan, Y.; McCallum, S. A.; Wang, C.; Garcia, A. E. J. *Mol. Biol.* **2007**, *368*, 1448–1457.
- (36) Granata, D.; Baftizadeh, F.; Habchi, J.; Galvagnion, C.; De Simone, A.; Camilloni, C.; Laio, A.; Vendruscolo, M. *Sci. Rep.* **2015**, *5*, No. 15449, DOI: [10.1038/srep15449](https://doi.org/10.1038/srep15449).

- (37) Krupa, P.; Huy, P. D. Q.; Li, M. S. J. *J. Chem. Phys.* **2019**, *151*, No. 055101, DOI: 10.1063/1.5093184.
- (38) Song, W.; Wang, Y.; Colletier, J.-P.; Yang, H.; Xu, Y. *Sci. Rep.* **2015**, *5*, No. 11024, DOI: 10.1038/srep11024.
- (39) Best, R. B.; Zheng, W.; Mittal, J. J. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124.
- (40) Chebaro, Y.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2009**, *113*, 7668–7675.
- (41) Rosenman, D. J.; Wang, C.; García, A. E. *J. Phys. Chem. B* **2016**, *120*, 259–277.
- (42) Okumura, H.; Itoh, S. G. *Molecules* **2022**, *27*, No. 2483.
- (43) Gremer, L.; Schölzel, D.; Schenk, C.; Reinartz, E.; Labahn, J.; Ravelli, R. B. G.; Tusche, M.; Lopez-Iglesias, C.; Hoyer, W.; Heise, H.; Willbold, D.; Schröder, G. F. *Science* **2017**, *358*, 116–119.
- (44) Yang, Y.; Arseni, D.; Zhang, W.; Huang, M.; Lövestam, S.; Schweighauser, M.; Kotecha, A.; Murzin, A. G.; Peak-Chew, S. Y.; Macdonald, J.; Lavenir, I.; Garringer, H. J.; Gelpi, E.; Newell, K. L.; Kovacs, G. G.; Vidal, R.; Ghetti, B.; Ryskeldi-Falcon, B.; Scheres, S. H. W.; Goedert, M.; et al. *Science* **2022**, *375*, 167–172.
- (45) Kim, M. O.; Nichols, S. E.; Wang, Y.; McCammon, J. A. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 235–246.
- (46) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (47) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1–2*, 19–25.
- (48) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185*, 604–613.
- (49) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, No. 014101, DOI: 10.1063/1.2408420.
- (50) Parrinello, M.; Rahman, A. *Phys. Rev. Lett.* **1980**, *45*, No. 1196.
- (51) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (52) Pietrucci, F.; Laio, A. *J. Chem. Theory Comput.* **2009**, *5*, 2197–2201.
- (53) Heinig, M.; Frishman, D. *Nucleic Acids Res.* **2004**, *32*, W500–W502.
- (54) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
- (55) Torrie, G.; Valleau, J. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (56) Bonomi, M.; Barducci, A.; Parrinello, M. *J. Comput. Chem.* **2009**, *30*, 1615–1621.
- (57) Marinova, V.; Salvalaglio, M. *J. Chem. Phys.* **2019**, *151*, No. 164115, DOI: 10.1063/1.5123498.
- (58) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- (59) Gowers, R.; Linke, M.; Barnoud, J.; Reddy, T.; Melo, M.; Seyler, S.; Domański, J.; Dotson, D.; Buchoux, S.; Kenney, I.; Beckstein, O. *MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations*, Proceedings of the Python in Science Conference; SciPy, 2016; pp 98–105.
- (60) Itoh, S. G.; Yagi-Utsumi, M.; Kato, K.; Okumura, H. *ACS Chem. Neurosci.* **2022**, *13*, 3139–3151.
- (61) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. *J. Comput. Chem.* **1995**, *16*, 273–284.
- (62) Roseman, M. A. *J. Mol. Biol.* **1988**, *200*, 513–522.
- (63) Villani, C. *Optimal Transport; Springer Berlin Heidelberg; Springer Berlin Heidelberg*, 2009.
- (64) Peyré, G.; Cuturi, M. *Found. Trends Mach. Learn.* **2019**, *11*, 355–607.
- (65) Murtagh, F.; Legendre, P. *J. Classif.* **2014**, *31*, 274–295.
- (66) Thorndike, R. L. *Psychometrika* **1953**, *18*, 267–276.
- (67) Carli, M.; Laio, A. *Mol. Phys.* **2021**, *119*, No. e1899323, DOI: 10.1080/00268976.2021.1899323.
- (68) Carli, M.; Rodriguez, A.; Laio, A.; Glielmo, A. *Mach. Learn.: Sci. Technol.* **2025**, *6*, No. 025052.
- (69) Ferreira, A.; Guo, R.; Marziano, I.; Salvalaglio, M. *J. Chem. Theory Comput.* **2025**, *21*, 12466–12480.
- (70) Barz, B.; Buell, A. K.; Nath, S. *Chem. Commun.* **2021**, *57*, 947–950.
- (71) Schäffler, M.; Wales, D. J.; Strodel, B. *Chem. Commun.* **2024**, *60*, 13574–13577.
- (72) Ball, K.; Phillips, A.; Wemmer, D.; Head-Gordon, T. *Biophys. J.* **2013**, *104*, 2714–2724.
- (73) Jokar, S.; Khazaei, S.; Gameshgoi, X. E.; Khafaji, M.; Yarani, B.; Sharifzadeh, M.; Beiki, D.; Bavi, O. *Alzheimer's Disease: Drug Discovery; Exon Publications; Exon Publications*, 2020; pp 51–68.
- (74) King, K. M.; Zaheer, H.; Brown, A. M. *Biophys. J.* **2025**, *124*, 2995–3007.
- (75) Khaled, M.; Rönnbäck, I.; Ilag, L. L.; Gräslund, A.; Strodel, B.; Österlund, N. *J. Am. Chem. Soc.* **2023**, *145*, 18340–18354.
- (76) Yan, Y.; Wang, C. *J. Mol. Biol.* **2006**, *364*, 853–862.
- (77) Lin, Y.; Im, H.; Diem, L. T.; Ham, S. *Biochem. Biophys. Res. Commun.* **2019**, *510*, 442–448.
- (78) Paul, A.; Samantray, S.; Anteghini, M.; Khaled, M.; Strodel, B. *Chem. Sci.* **2021**, *12*, 6652–6669.
- (79) Marinova, V.; Dodd, L.; Lee, S.-J.; Wood, G. P. F.; Marziano, I.; Salvalaglio, M. *J. Chem. Inf. Model.* **2021**, *61*, 2263–2273.
- (80) Monti, S.; Tamayo, P.; Mesirov, J.; Golub, T. *Mach. Learn.* **2003**, *52*, 91–118.
- (81) Rodriguez, A.; Laio, A. *Science* **2014**, *344*, 1492–1496.
- (82) Rousseeuw, P. J. *J. Comput. Appl. Math.* **1987**, *20*, 53–65.
- (83) Kollmer, M.; Close, W.; Funk, L.; Rasmussen, J.; Bsoul, A.; Schierhorn, A.; Schmidt, M.; Sigurdson, C. J.; Jucker, M.; Fändrich, M. *Nat. Commun.* **2019**, *10*, No. 4760, DOI: 10.1038/s41467-019-12683-8.
- (84) Baek, Y.; Lee, M. *Biochem. Soc. Trans.* **2024**, *52*, 1631–1646.
- (85) Lee, M.; Yau, W.-M.; Louis, J. M.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120*, No. e2218831120, DOI: 10.1073/pnas.2218831120.
- (86) Tarus, B.; Tran, T. T.; Nasica-Labouze, J.; Sterpone, F.; Nguyen, P. H.; Derreumaux, P. *J. Phys. Chem. B* **2015**, *119*, 10478–10487.
- (87) Mehrazma, B.; Rauk, A. *J. Phys. Chem. A* **2019**, *123*, 4658–4670.
- (88) Brown, A.; Bevan, D. *Biophys. J.* **2016**, *111*, 937–949.
- (89) Barz, B.; Liao, Q.; Strodel, B. *J. Am. Chem. Soc.* **2018**, *140*, 319–327.



CAS INSIGHTS™

## EXPLORE THE INNOVATIONS SHAPING TOMORROW

Discover the latest scientific research and trends with CAS Insights. Subscribe for email updates on new articles, reports, and webinars at the intersection of science and innovation.

[Subscribe today](#)

**CAS**  
A Division of the  
American Chemical Society