JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

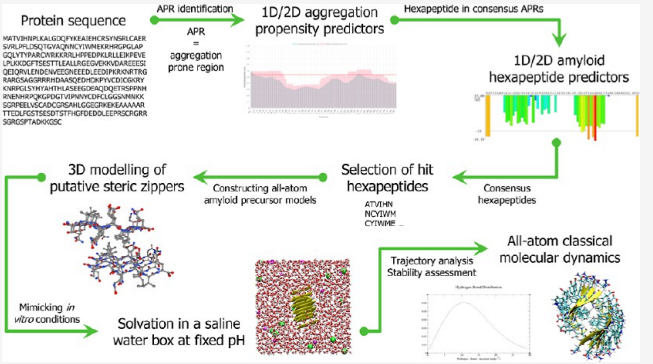
pubs.acs.org/jcim Article

Identification of Amyloid Regions and Mechanisms from Sequence-Based Modeling and Molecular Dynamics Simulation: A Case Study of the Intrinsically Disordered Protein DPF3

Julien Mignon,* Tanguy Leyder, Hugo Bâlon, Antonio Monari, Denis Mottet, and Catherine Michaux

Cite This: J. Chem. Inf. Model. 2026, 66, 5471−5494 Read Online

ABSTRACT: Amyloids refer to a diverse group of highly ordered proteinaceous structures conserved across all of the domains of life that not only are involved in severe proteinopathies but also serve as unique functional platforms. They are notably organized around a cross-β sheet core arranged into a dry steric zipper. Identifying regions promoting such type of assembly requires investigating, usually at the atomistic level, their structural and morphological properties. Although more amyloid structures are being solved and described by means of high-resolution experimental techniques, most amyloidogenic systems remain experimentally elusive or uncharacterized. To overcome such limitations, we propose a multilevel sequence-based computational approach taking advantage of aggregation-oriented predictors and modelers coupled with all-atom equilibrium molecular dynamics (MD) simulations. As an example of this approach, we introduced our protein of interest, the intrinsically disordered zinc finger DPF3a, to the pipeline, thus identifying a hit hexapeptide (^{52}NCYIWM^{57}) exhibiting remarkable pro-amyloid features. Indeed, this peptide was shown to stabilize different steric zipper topologies, spontaneously self-assemble into β-sheeted oligomers, and to act as an amyloid seed enabling the conversion and addition of β-sheet multimers for fibril elongation, which echoes the pathological repertoire of DPF3a. In comparison, simulations performed on a peptide sequence consensually predicted as nonamyloid (^{33}AERSVR^{38}) congruently underscored its inability to maintain a steric zipper configuration and nucleate into oligomers enriched in β character. Using only the sequence of a target protein as starting input, our methodology noteworthily proved to be reliable for detecting aggregation-prone regions, assessing their amyloidogenicity, and elucidating their fibrillation mechanisms while holding promise for the design of new anti-amyloid drugs and amyloid-inspired biomaterials.



1. INTRODUCTION

Ordered and intrinsically disordered proteins are susceptible to undergo improper or out-of-pathway folding under cellular stress or dysregulation, leading to the formation of insoluble aggregated species associated with the development and progression of proteinopathies, such as Alzheimer's and Parkinson's diseases.[1,2] The latter are characterized by the misfolding and aggregation of proteins in several parts of the brain.[3,4] One of the most remarkable hallmarks in such neurodegenerative disorders is the formation and accumulation in neurons of highly structured, fibrous, and insoluble supramolecular assemblies, called amyloid fibrils.[5] Apart from neurodegeneration, such proteinaceous filaments are also known to form extracellular deposits in diverse tissues, such as heart, kidney, liver, or muscles, building amyloid material up over time and subsequently leading to either systemic or localized amyloidosis due to organ failure.[6]

However, many organisms have also beneficially taken advantage of such stable and symmetrical fibrillar self-assemblies over the course of evolution. Referred to as functional amyloids in opposition to pathogenic oligomers and fibrils, they are notably exploited by bacteria or fungi for adhering to surfaces, as well for establishing chemoresistant biofilms, which represent a resistance strategy deployed against antibiotic and antifungal drugs.[7,8] Conversely, amphibians have developed barriers to counteract bacterial infections by secreting on their skin fibril-forming antimicrobial peptides.[9,10] Among biomaterials, silk proteins produced by insects or spiders are well-known for sharing structural features and mechanical properties with amyloid filaments while avoiding the generation of cytotoxic intermediates or deposits.[11]

Notwithstanding their pathogenic and functional diversity, amyloid fibrils commonly display distinctive molecular and

Received: March 27, 2026
Revised: April 8, 2026
Accepted: April 10, 2026
Published: April 16, 2026

architectural characteristics compared to other types of protein aggregates or nanoscale assemblies. Their spine is most notably built around a typical cross-β core, featuring two complementary β-sheets perpendicularly stacked to the fibril axis with interstrand and intersheet distances of around 4.8 and 10.0 Å, respectively.[12,13] Such configuration allows for the clustering and interdigitation of hydrophobic side chains within a tight and anhydrous steric zipper that diversifies into several symmetry classes. Depending on the direction of the peptide chain (up or down), the orientation of amino acid side chains (face or back), and the organization of β-strands (parallel or antiparallel), as well as the face composition of contacting β-sheets (anti- or equifacial), steric zippers allocate into eight topological combinations (classes 1 to 8) that have been experimentally evidenced by X-ray diffraction or electron microscopy. In this context, the face refers to the preponderantly hydrophobic side enclosed within the zipper.[14,15] Nevertheless, among the 15 mathematically possible homosteric zipper groups, other putative configurations have been computationally explored, such as class 1- and class 3-like of the PHF6 peptide found in the sequence of the tau protein.[16,17]

To add to the complexity, the two β-sheets can consist of distinct polypeptides, resulting in the creation of heterosteric zippers stabilized through heterotypic interactions. In the past decade, this has become particularly true for coaggregating proteins, forming heterogeneous fibrillar species containing molecules with distinct sequences.[18,19] For example, *in vivo*, *in vitro*, and *in silico* studies have compellingly demonstrated that amyloid β (Aβ) and tau have a synergistic relationship in Alzheimer's disease through cross-interaction and hetero-oligomerization.[20] Furthermore, it should be pointed out that noncanonical cross-α fibrils, undiscovered since 2017, exist. These reversible filamentous architectures most interestingly endorse plural biological functions, such as biofilm formation or antimicrobial activity, while expanding the structure−function paradigm of the proteinaceous fibrillated state.[21]

The processes of zipper packing and fibrillation are mediated by aggregation prone regions (APRs). Such stretches can be located in intrinsically disordered regions and possess common sequence motifs, i.e., high hydrophobicity, low local charge, and favorable β-sheet propensity, as encountered for (iso)-leucine and valine,[22] that promote protein self-assembly while burying into a dry interface.[23,24] Additional factors, such as side chain−side chain H bonds, salt bridges, and π stacking, as well CH−π and ion−π interactions, should be considered since they are directional organizers and stabilizers of steric zippers and fibrillar structures.[25] Taken together, these effects funnel the conformational space of amyloid fibrils to highly thermodynamically stable energy minima. Nevertheless, the free energy landscape of the amyloidogenic pathway remains complex, as it comprises a great diversity of conformational states, such as partly folded states, amorphous phase, oligomers, and protofibrils, which necessitate to cross inter-intermediate energy barriers to form fibrils over the course of protein aggregation.[13,26] Fibril growth is indeed a nucleation-dependent and multistep polymerization process characterized by several intermediates along the conformation coordinate. Initially, the protein needs to adopt an amyloidogenic conformation through structural rearrangement and APR exposure, which in turns nucleates into oligomers elongating into protofibrils, to finally assemble into mature fibrils.[27] The primary nucleation event is the rate-limiting step and is determined by both intrinsic and external factors as well as the properties of the water hydration shell at the protein or peptide surface. While intermolecular contacts are thermodynamically favored due to the hydrophobic effect, they are restricted by large desolvation energy barriers.

Therefore, gaining insight into the molecular mechanism of protein-specific self-assembly and identifying the APRs that initiate β-sheet conversion is crucial for targeting amyloidogenic proteins in proteinopathies through the development of anti-amyloid strategies that can interfere with such regions, sequester molecules into off-pathway oligomeric species, or disassemble amyloid precursors.[28] Such prospects are not limited to therapeutical applications. Evaluating the aggregation and amyloidogenic propensity of nature-based or synthetically tuned peptides for the design of amyloid-inspired biomaterials, nanoplatforms, or antimicrobial agents is of great interest and relevance nowadays.[29]

However, some biomolecular or artificial systems are either continuously overlooked in the literature, still uncharacterized, or experimentally challenging in such a sense that the description of their structural and mechanistic properties eludes traditional biophysical or high-resolution techniques. In this context, we have worked during the last five years on a whimsical intrinsically disordered protein, known as the double PHD fingers 3 (DPF3) protein, existing as two isoforms (DPF3b and DPF3a). DPF3 is a zinc finger epigenetic regulator involved not only in chromatin remodeling and mitosis but also in cancer and neurodegeneration.[30−33] The isoforms are identical up to the 292nd position but differentiate themselves by their domain composition at the C-terminus. Whereas DPF3b is 378 residues long and contains a double PHD zinc finger domain that reads modified histone tails, DPF3a is truncated to 357 residues and possesses a disordered C-terminal region, the function of which remains unclear. Albeit unveiling their environmental pro-amyloid susceptibility and amyloid-related optical fingerprints, no structure of DPF3 fibrils has been solved up to now, curtailing our molecular-level knowledge on the aggregation of its two splicing variants and on which of their domains compose the amyloid core in the fibrillated state.[33−38] To overcome these limitations, we developed a sequence-based multilevel computational pipeline for the identification of APRs and amyloid peptides, the construction of steric zippers, and the exploration of their self-assembly pathways. This is achieved by exploiting freely available prediction and modeling tools coupled with all-atom classical molecular dynamics (MD) simulations. By submitting the DPF3a protein to the test, the methodology behind our approach, as well as its reliability for discovering new amyloid-driving peptides and domains in protein sequences, is hereafter described and discussed.

2. METHODS

2.1. Prediction of Aggregation Prone Regions and Selection of Hit Hexapeptides

The protein sequence of the human DPF3a isoform was retrieved from the UniProt database (UniProt entry: Q92784-2) and submitted throughout the following aggregation-oriented predictors for the detection and selection of APRs (step i): AGGRESCAN,[39] Agmata,[40] AmyloGram (self-assembly probability cutoff of 0.50),[41] ArchCandy-2.0 (score threshold of 0.40),[42] CamSol (condition-defined pH value, here pH 8.0),[43] Cross-Beta RF (score threshold of 0.54),[44] Pafig,[45] TANGO,[46] and WALTZ.[47] To define the thresholds, we drew inspiration from previously developed consensus-based methods for

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

aggregation predictions, considering the hit overlap of at least $n/2$ (rounded down) out of n algorithms.[48,49] Among the output APRs, consensus was therefore established out of at least four occurrences across the nine independent predictors for the same sequence stretch with a discrepancy window of three positions at each end. Along the whole protein sequence, with a particular focus within the previously defined consensus APRs, pro-amyloid hexapeptides, corresponding to the minimum length for a steric zipper rearrangement, were identified using the following amyloid-oriented predictors (step ii): ANuPP (score threshold of 0.45),[50] Budapest Amyloid Predictor,[51] FiSH Amyloid (full prediction mode with a score threshold of 0.19),[52] FoldAmyloid (score threshold of 21.4),[53] Generalized Aggregation Proneness,[54] MetAmyl (best global accuracy),[55] and ZipperDB (score threshold of −23 kcal/mol).[56] For the selection of hit amyloid hexapeptides, consensus was determined out of at least three occurrences across the seven independent predictors for the same peptide segment, with a discrepancy window of one position at each end.

2.2. Model Construction of Putative Steric Zippers

One hit amyloid hexapeptide along with one hexapeptide consensually predicted as nonamyloid for negative control was enforced into putative homosteric zippers, containing five copies of each peptide in each of the two facing β-sheets for a total of 10 stacked chains (step iii). For that purpose, three experimentally validated modelers, i.e., CORDAX, FibPredictor (FibPred), and PATH, were used. The most stable first ranked model output by each of them was selected for subsequent all-atom MD simulations. CORDAX is a structure-based machine learning approach that returns, among other template zippers, the most stable model after energy minimization in the empirical FoldX force field.[57,58] *De novo* modeling software FibPred relies on geometric transformations from a β-fibril backbone and the statistical function Amb_3b for energy scoring.[59,60] The following modeling parameters were used: 30 translation vectors for the number of models, 10 top models for ranking, as well as all axis rotations, a distance variation and minimum distance between the sheets of 8.0 and 3.5 Å, respectively, and an angle variation of 15° between the sheets for the generation of zipper symmetries. PATH consists in a threading-based machine learning algorithm that uses solved structures of peptide steric zippers corresponding to the eight experimentally evidenced topology classes as templates and ranks the models quality according to the statistical potential DOPE as the scoring method.[61,62] For further positive control, the amyloid PHF6 ^{306}VQIVYK^{311} hexapeptide of the tau protein was also enforced into a homosteric zipper with the CORDAX modeler.

2.3. System Preparation for Molecular Dynamics Simulations

Prior to MD simulation in explicit solvent, the generated models were first properly protonated at the level of the backbone amide and side chain atoms using the PDB Reader & Manipulator tool available in the CHARMM-GUI suite.[63,64] Within the same module, each peptide chain was capped by acetylating and amidating its N- and C-terminus, respectively, to avoid the introduction of a nonrealistic charge effect on peptide ends. This step is required only if the hexapeptide is neither N- nor C-terminal. By exploiting the CHARMM-GUI Solution Builder module, the steric zipper model was centered in a cubic water box in which a 10 Å buffer of water molecules was added from each edge.[65] Depending on the amino acid composition of the modeled zipper, the required minimal amount of Na^+ or Cl^- ions was added to ensure system neutralization. To account for experimental conditions, additional ions were randomly and accordingly introduced with respect to the simulation box volume to achieve the physiologically relevant NaCl concentration of 150 mM. Subsequently, the extended random coil structure of the corresponding hexapeptide sequences from DPF3 and tau was modeled, capped, and protonated using the Builder plugin in PyMOL.[66] Ten copies of the generated peptide conformer were randomly placed at a minimal 10 Å distance apart from one another and subsequently solvated within a cubic water box with 50 Å sides using CHARMM-GUI Multicomponent Assembler.[67] Finally, the central structure of the most populated cluster for one of the relaxed steric zippers obtained after trajectory dimensionality reduction was introduced in the free hexapeptide system and surrounded by water molecules in a cubic box with 70 Å sides. Trajectory clustering was performed beforehand using a root-mean-square deviation (RMSD) cutoff of 1.5 Å between the nearest neighbors with a time step of 1 ns. As previously described, charge neutralization and near-physiological salt concentration were guaranteed by adding the adequate amount of Na^+ and Cl^- ions.

2.4. Unbiased All-Atom Equilibrium Molecular Dynamics and Trajectory Analysis

All-atom classical MD simulations were carried out for all the systems, i.e., steric zippers (step iv), precursor-free monomers (step v), and seeded monomers (step vi), in independent triplicates with the GROMACS 2023.1 suite implemented with GPU (CUDA 11.7.0) accelerators.[68,69] Given their reported performance for simulating aggregation processes in proteins, the peptides were modeled with the atomistic CHARMM36m force field used in conjunction with the TIP3P water model modified for CHARMM.[70−72] Integration of the Newton's equations of motion was achieved with a time step of 4 fs upon relying on hydrogen mass repartitioning, directly implemented in CHARMM-GUI and applied on the initial topologies.[73,74] Furthermore, bonds involving hydrogen atoms were constrained to their equilibrium length using the LINCS algorithm.[75] The modeled systems were first minimized for a maximum of 5000 steps using the steepest descent algorithm.[76] Using the leapfrog integrator for integrating the equations of motion, the simulation boxes were afterward thermalized, equilibrated, and propagated in the isothermal and isobaric (NPT) ensemble.[77] A pressure of 1 bar and a temperature of 300 K were respectively enforced by a C-rescale and a modified Berendsen thermostat (V-rescale).[78,79] Short-range nonbonded interactions were computed with a distance cutoff of 1.2 nm, whereas particle mesh Ewald summation was used for long-range Coulombic interactions.[80] Production MD triplicates were propagated under periodic boundary conditions in all three dimensions for a duration of 1 μs each. Along the simulations, the atom coordinates and energies were recorded every 100 ps. All simulated parameters are summarized in Tables S1 and S2 for DPF3 and tau systems, respectively.

Modules directly invoked from the GROMACS suite were conjunctively used with in-house built Python scripts for analyzing and clustering the trajectories. For each system, time-dependent and per-residue variables and their associated probability density functions (PDF) were computed and averaged across the replicates. RMSD was calculated from the equilibrated structure, while root-mean-square fluctuation (RMSF) was determined for each amino acid by taking the time-averaged structure as a reference. Over the replicates, a total of 24,000 frames were clustered by defining 1.5 and 5.0 Å RMSD cutoffs between the nearest neighbors for the zippers and precursor-free monomers, respectively. For the generation of reduced free energy surfaces (FES), principal component analysis (PCA) was carried out on the most representative trajectory selected after clustering using the RMSD of the backbone atoms for building the covariance matrix. Along the two first principal components capturing the essential motions of the sampled systems (PC1 and 2), probabilities were computed and distributed in a hundred of basins with the kernel density estimation, and free energy variation (in J/mol) was calculated as follows: $\Delta G(\mathrm{PC}) = -RT[\ln P(\mathrm{PC}) - \ln P_{\max}]$, where R is the ideal gas constant, T is the temperature (in K), P is the probability distribution, and $P_{\max}$ is the probability distribution maximum. To account for intra- and intermolecular H bonds, 3.5 Å donor−acceptor maximum distance and 30° angle cutoffs were applied. Regarding secondary structure classes, they were defined using the STRIDE algorithm implemented in the Timeline VMD plugin: turns (predominantly β type, $i-i+3$ turn), extended β-sheet, isolated β-bridge (usually $i-i+>6$), α-helix ($i-i+4$ helix), 3_{10}-helix ($i-i+3$ helix), π-helix (distorted or bulged $i-i+5$ helix), and coil.[81]

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

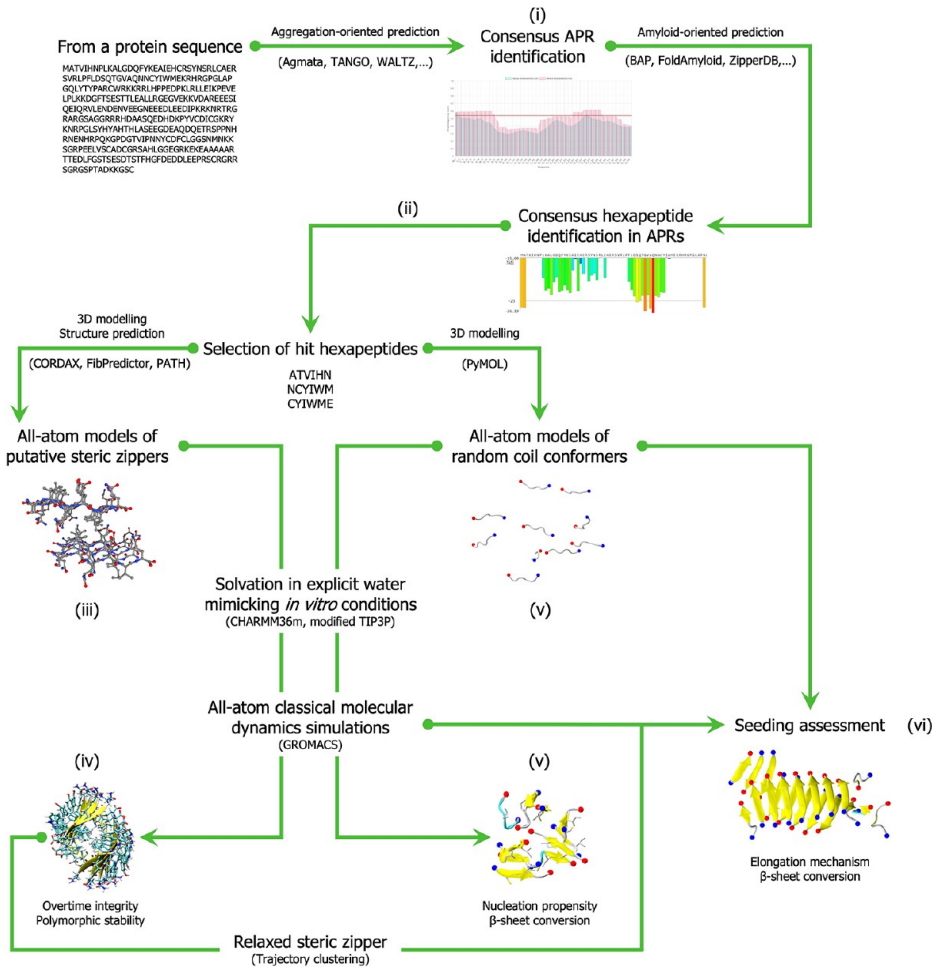


Figure 1. Schematized overview of the computational protocol for identifying amyloid regions and peptides. The presence of APRs is scanned along the query protein or peptide sequence (i) in which amyloid hexapeptides are selected upon reaching consensus among predictors (ii). Hit peptides are enforced and modeled into cross-β steric zipper configurations using amyloid-trained algorithms (iii) before relaxation in explicit water through all-atom equilibrium MD to assess their stability and polymorphism susceptibility (iv). Alongside zipper persistence, nucleation behavior is monitored by modeling the peptides into a random coil conformation and MD simulation (v). Finally, previously relaxed zippers are extracted and solvated around unfolded peptide chains to get insights into the seeding and elongation processes by carrying out MD simulations (vi).

Trajectory visualization and snapshot rendering were performed with the VMD molecular graphic program.[82]

3. RESULTS AND DISCUSSION

3.1. Overview of the Computational Pipeline

After scouring the specialized literature, we compiled a panel of sequence-based and aggregation-oriented algorithms to enhance the robustness of predictions through cross-validation over existent predictors. In that respect, the proposed methodology takes advantage of a large array of tools that are freely available online and accessible through user-friendly Web servers, software, or shared codes. In addition, it is very adaptative to any type of protein or peptide system, especially those eluding *in vitro* characterization of their structural components, as it only requires the sequence of the biomolecular target of interest as the starting input. Therefore, every output will be derived from the primary structure as the initial entry information. The computational pipeline revolves around six main successive steps detailed hereafter (Figure 1):

(i) Identification of APRs by reaching an arbitrary $n/2$ (rounded down) consensus among selected n aggregation-oriented predictors.

(ii) Computation of pro-amyloid scores and identification of amyloid regions and, more specifically, hexapeptides within previously defined APRs upon reaching an arbitrary $n/2$ (rounded down) consensus among selected n amyloid-oriented predictors.

(iii) Modeling of hit hexapeptides into putative three-dimensional and atomistic steric zipper organizations, composed of two mating and finite β-sheets.

(iv) Relaxation and integrity monitoring of modeled steric zipper precursors in an explicit solvent and saline environment mimicking *in vitro* or pathophysiological conditions by all-atom classical MD at the equilibrium.

(v) Simulation of multiple copies (e.g., 10) of the free monomer in solution starting from random coil conformers of the corresponding hexapeptide to assess their inherent propensity to convert into β-sheeted oligomers (nucleation step).

(vi) One-system simulation of a previously relaxed amyloid steric zipper (at step iv) in the presence of multiple copies (e.g., 10) of the free monomer of the corresponding hexapeptide starting in random coil conformation to assess the seeding ability of the precursor (elongation step).

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

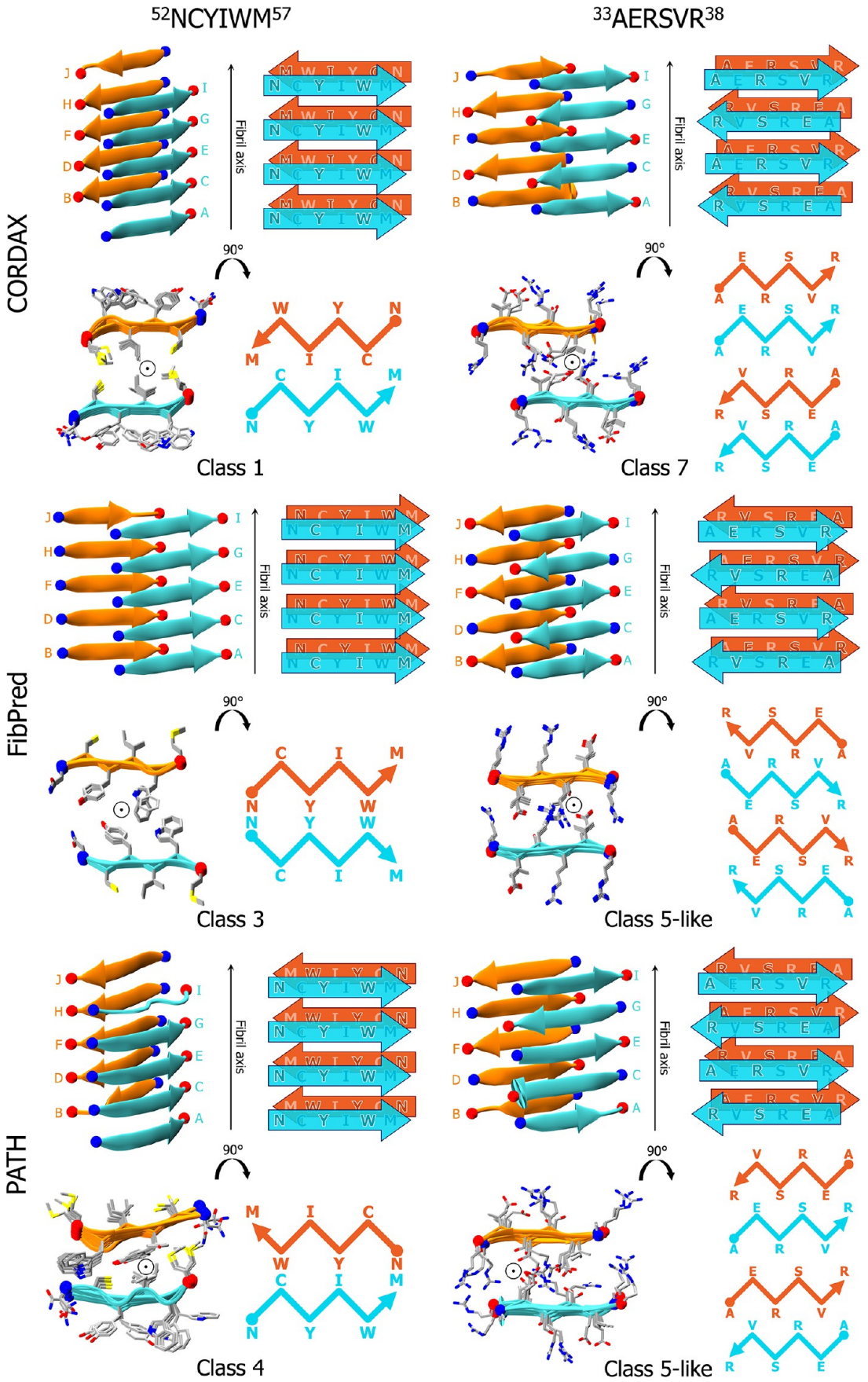


Figure 2. Starting models of the ^{52}NCYIWM^{57} and ^{33}AERSVR^{38} hexapeptides arranged into steric zippers with the selected algorithms. On each panel, peptide chains (annotated from A to J) are displayed in cartoon representation with the N- and C-termini respectively pinpointed as blue and red spheres, and the fibril axis depicted by a black arrow. For every combination of peptide−modeler, the side view (top) is compared to the

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

Figure 2. continued

top view (bottom) of the generated steric zipper, the front and rear β-sheets of which are colored in cyan and orange, respectively. The side view, perpendicular to the fibril axis, shows the arrangement of β-sheets and strands within each, whereas the top view, down the fibril axis, evidences the orientation of amino acid side chains, highlighted in licorice representation. Additionally, the organization of steric zipper classes is schematically illustrated for each corresponding pair of peptide−modeler from the top and side views to better appreciate the composition of the zipper core and its solvent-exposed surface.

3.2. Identification of Aggregation Prone Regions from Sequence-Based Predictions

Upon introducing the DPF3a protein primary structure to our computational pipeline, we therefore searched for regions and peptides consensually predicted as aggregation prone and, after refinement, pro-amyloid (Table S3). Based on our previously defined consensus thresholds, one hit amyloid hexapeptide, corresponding to ^{52}NCYIWM^{57}, stood out among others and was selected for subsequent three-dimensional modeling into putative steric zippers. For comparison purpose and strengthening the relevance of our selection process, the ^{33}AERSVR^{38} peptide was additionally picked out, as the sequence stretch it encompasses is unambiguously assumed to be nonamyloid for most predictors (Table S3). Interestingly, those two hexapeptides are found within the N-terminal region of the DPF3a protein, a domain that is commonly shared with its other isoform, known as DPF3b. As such, results discussed hereafter are transferrable to the two splicing variants of DPF3 and will be informative about their overall amyloid pathway and susceptibility that were experimentally ascertained in the past few years.[33−38]

3.3. From Sequence to 3D Structures: Modeling of Putative Steric Zippers

For assessing the (non)amyloidogenicity of identified hexapeptides, we took advantage of recently developed tools which are specifically designed for modeling cross-β assemblies and performed well on recognized amyloid proteins.[58,60,62] Three independent computational tools were used, taking hexapeptide sequences as input: (i) the structure-based machine learning software CORDAX, (ii) *de novo* modeling based on an energy scoring method FibPredictor (FibPred), and (iii) the threading-based machine learning algorithm PATH. For each peptide−modeler combination, the first top-ranked structure was retrieved and visualized (Figure 2).

Confirming that ^{52}NCYIWM^{57} is pro-amyloid, CORDAX predicts that the most stable arrangement is achieved into a class 1 steric zipper, i.e., parallel, up−up, and face-to-face in-register β-strands, for which the face is defined according to the isoleucine (Ile) side chain inward orientation (Figure 2, top left). Even prior to MD relaxation, it can already be pinpointed that three major effects are likely responsible for promoting a face-to-face class 1 zipper: (i) hydrophobic packing of Ile aliphatic chains at the dry interface, (ii) a H-bonded asparagine (Asn) N−H···O ladder at the N-terminal edge of the rear sheet (Figure S1A), and (iii) sandwich-like $\pi-\pi$ stacking of surface-exposed tyrosyl (Tyr) and tryptophanyl (Trp) aromatic rings (Figure S1B). Indeed, Asn ladders[15,83] and aromatic π stacking[84,85] are well-known stabilizers of amyloid fibrils, participating in their self-assembly and defining their directionality.

Comparatively, FibPred assigns a class 3 symmetry to ^{52}NCYIWM^{57}, consisting of parallel, up−down, and face-to-face in-register β-strands, which, instead of Ile, encloses Tyr and Trp within the dry interface (Figure 2, middle left). With respect to the front sheet, the face corresponds here to the one comprising Tyr and Trp side chains, while the up−down nomenclature indicates that the rear sheet is flipped 180°. One of the outcomes of such topology implies that the front and rear β-sheets are parallel to each other, which, in this case, can hinder the formation of a tight zipper due to the inherent steric hindrance of aromatic rings. Nevertheless, the FibPred model presents a slight increase in intersheet offset and spacing that seemingly better accommodates the volume occupied by aromatic residues in the zipper while allowing CH−π interactions between Trp side chains (Figure S1E).[86] Besides those considerations, an overall stabilization in each β-sheet is achieved, as for the CORDAX class 1 model, through the combination of Asn H bonding (Figure S1C) at the N-termini with intrasheet Tyr and Trp sandwich $\pi-\pi$ stacking (Figure S1D).

The steric zipper generated with PATH exhibits a class 4 geometry, that is, parallel, up−down, and face-to-back in-register β-strands, giving rise to a mixed dry interface with the face of the front sheet seeing the back of the rear sheet (Figure 2, bottom left). Although weaker than π stacking, CH−π interactions between facing Ile and Tyr residues nonetheless contribute to the zipper core stabilization thanks to the Tyr planar side-to-side alignment (Figure S1F).[87] In addition, the parallel orientation can still enable, to some extent, the formation of an Asn ladder and $\pi-\pi$ stacked Tyr/Trp residues (Figure S1G). The steric zipper predicted by PATH appears more unstable, notably due to the insufficient intersheet spacing with respect to the size of embedded aromatic rings. However, such unfavorable steric clashes could be relaxed by MD. Furthermore, interactions between Tyr residues predominantly occur at the zipper surface and are partly driven by the polarity of the hydroxyl moieties.[88] By comparing the tyrosyl orientation between the models, such a preference still advocates in favor of the class 1 organization, as classes 3 and 4 respectively confine all or half of Tyr side chains in the hydrophobic core. Overall, the parallel packing of strands and antifacial sheets observed among the models could have been anticipated, as the presence of an Asn residue in ^{52}NCYIWM^{57} permits their alignment in a ladder-like topology, promoting the formation of parallel β-sheets while enhancing the assembly stability.[89]

Regarding ^{33}AERSVR^{38}, CORDAX expectedly considers the peptide as nonamyloid, but, when enforcing it into a steric zipper, assigns a likely class 7 topology, meaning antiparallel, up−up, face equals back, in-register β-strands, with the front and rear sheets being equifacial and parallel to one another (Figure 2, top right). Indeed, the face corresponds to the back, i.e., equifacial, because each strand is alternately rotated 180° around the fibril axis for antiparallel stacking, thus exhibiting dry and wet interfaces of the same composition (Figure S2A). Interestingly, the two β-sheets are also offset, which may reduce steric hindrance and allow for a better solvation of C-terminal arginine (Arg) residues. Compared to ^{52}NCYIWM^{57}, the antiparallel orientation avoids unfavorable electrostatic

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

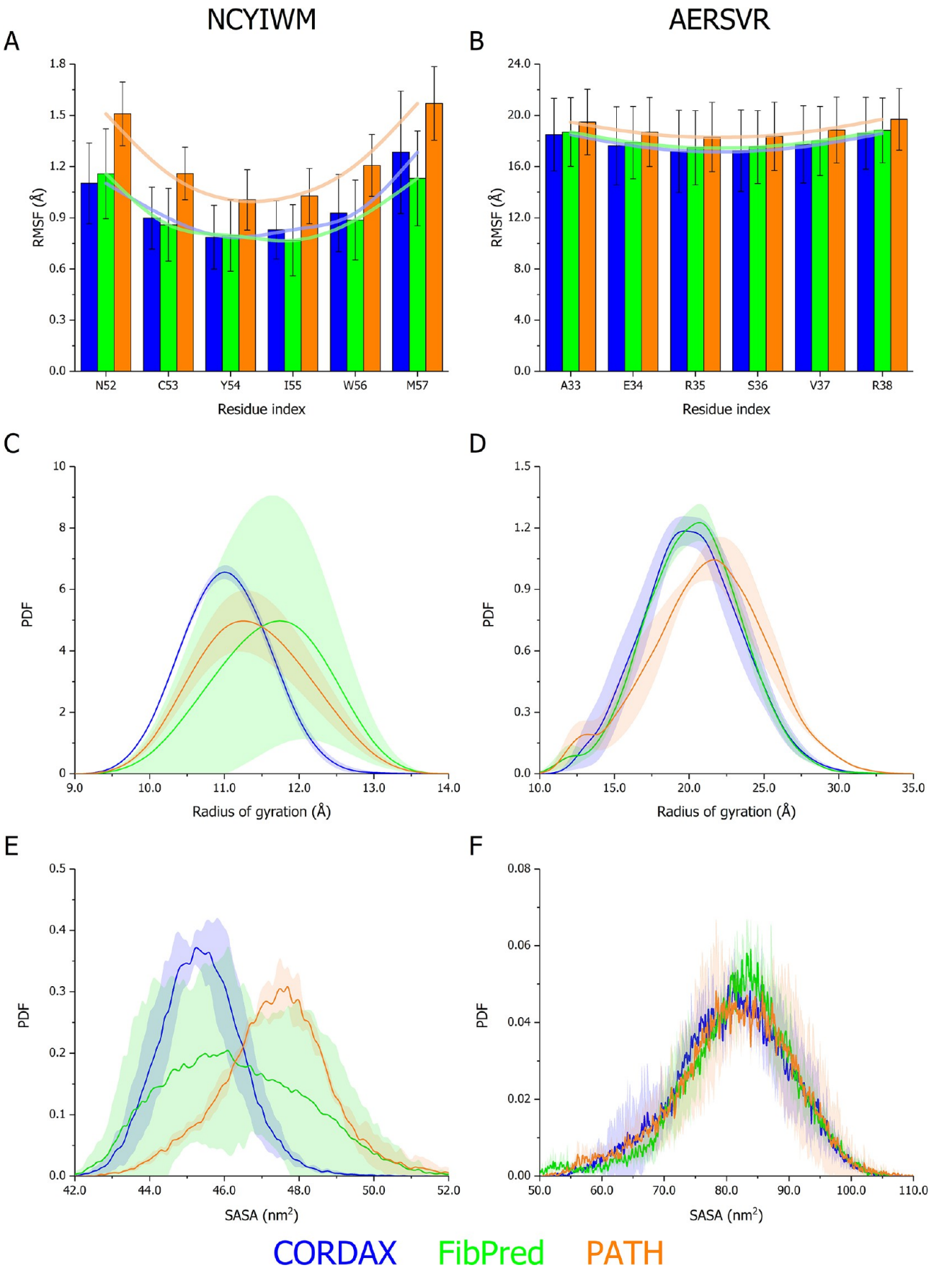


Figure 3. Flexibility, compaction, and hydration properties of ^{52}NCYIWM^{57} and ^{33}AERSVR^{38} steric zippers. (A, B) Time-averaged backbone RMSF, as well as probability density function (PDF) of (C, D) radius of gyration (R_g) and (E) solvent accessible surface area (SASA) of CORDAX (blue), FibPred (green), and PATH (orange) models of NCYIWM (A, C, E) and AERSVR (B, D, F) peptides simulated for 1 μs. On the RMSF

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

Figure 3. continued

panels (A, B), per-residue values have been averaged over the 10 hexapeptide chains constituting the initial steric zipper, standard deviations between triplicates are depicted as error bars, and global trends along the peptide sequence as curved lines in the model-associated color. On the R_g and SASA PDF panels (C−F), curves correspond to the average of triplicates with the standard deviation represented as a trace (shaded area) in the model-associated color.

repulsion between packed Arg side chains. Furthermore, in such configuration, the mixed interface partly stabilizes the zipper core via H bonds involving Arg, glutamate (Glu), and serine (Ser) residues within and between the facing β-sheets (Figure S2B). However, exclusion of valine (Val) side chains from hydrophobic packing might destabilize the assembly.

Conversely, FibPred (Figure 2, middle right) and PATH (Figure 2, bottom right) both predict the hexapeptide to adopt a class 5-like symmetry. However, while the class 5 steric zipper is typically composed of antiparallel, up equals down, face-to-face, out-of-register β-strands with antifacial antiparallel sheets, the models instead display a face equals back topology, in-register strands, and equifacial antiparallel sheets, resulting in identical surfaces but an alternating core composition of Ala−Arg−Val and Glu−Ser−Arg (Figures S2C, E). In that respect, the models comprise symmetry elements from classes 5, 7, and 8. According to a recent review, such orientation even matches a new class 8, diverging from previous classifications and defined with antiparallel β-sheets.[15] The FibPred structure is more or less the same as that previewed by PATH, although the PATH zipper appears to be slightly more compact. Similar to CORDAX class 7, stabilization is presumably mediated through intra- and intersheet H bonds and salt bridges involving Arg, Glu, and/or Ser amino acids (Figures S2D, F).

While exhibiting a certain variety of class patterns among the predictors, ^{52}NCYIWM^{57} seems to preferentially adopt an antifacial parallel orientation, which enhances the zipper stability through the addition of several effects, such as hydrophobic packing of aliphatic chains, Asn H-bonded ladders, $\pi-\pi$ stacking of aromatic rings, and CH$-\pi$ interactions. Comparatively, all modelers agree that the ^{33}AERSVR^{38} peptide fosters an antiparallel configuration to reduce electrostatic repulsion between Arg side chains at the cost of partial burial of Val residues. Moreover, the resulting mixed zipper cores are characterized by intra- and/or intersheet salt bridging between charged residues, which has also been reported for murine serum amyloid A, promoting the cohesion between protofilaments in mature fibrils,[25] or for the Aβ16-22 construct known to adopt a preferential antiparallel orientation driven by charge attraction between K16 and E22 residues.[90] In light of such observations, one can already anticipate why ^{52}NCYIWM^{57} was consensually predicted as pro-amyloid using sequence-based predictors, whereas ^{33}AERSVR^{38} failed due to its polar amino acid composition and dominating electrostatic repulsion that might preclude the persistence of a steric zipper conformation. These considerations were further examined through structural relaxation of modeled zippers by all-atom classical MD.

3.4. Simulation of Steric Zippers for an Amyloid Hexapeptide

Since the models generated with the three algorithms are heuristic and static structures that do not account for amino acid flexibility and, as underlined for several classes, can introduce unfavorable contacts, each steric zipper was simulated for 1 μs in an explicit solvation water box using near-physiological conditions in three independent replicates. First, the class 1, 3, and 4 topologies of the ^{52}NCYIWM^{57} hexapeptide were considered and their evolution along the trajectories is hereafter discussed.

Regardless of the steric zipper symmetry, triplicates are convergent and present a low root-mean-square (RMSD) deviation, indicating that the overall integrity of the zipper structure is maintained over time (Figure S3A). Compared to the CORDAX model, the FibPred and PATH structures nonetheless display higher variation between replicates. This is exacerbated for the PATH zipper, the conformation of which becomes more heterogeneous from 600 ns onward with a higher RMSD variability. Such RMSD increase originates from the flexibility of one of the chains, presumably localized at the edges of the structure; for example, chain I was lacking a β-sheeted fold in the starting model. The higher flexibility of class 4 is indeed evidenced by a global increase in RMSF across all residues with respect to the other two topologies (Figure 3A). Nevertheless, RMSF values are consistent for all the zippers in describing a more pronounced rigidity of the amino acids composing the core of the β-sheets (C53, Y54, I55, and W56). This also suggests that the Asn ladder should be considered partly dynamic rather than fully static. Such trends are coherently observed across the 10 chains constituting the amyloid structure (Figure S4A−C), also highlighting Y54 and I55 as the less flexible residues and the core peptides (chains C, E, G, D, F, and H) of each sheet as the most rigid.

Compaction (Figure S3C) and hydration (Figure S3E) parameters are in agreement with the RMSD and RMSF tendencies, respectively showing that the radius of gyration (R_g) and SASA exhibit only small transient variations. These small changes cannot be assigned to a global rearrangement of the supramolecular structure of the assembly and support that none of the peptide chains disassemble from the steric zipper along the total simulation span. While this is particularly true for class 1, exhibiting barely any visible deviation between replicates, FibPred curves are characterized by stronger variability, as well as higher plateaus, leading to broader R_g and SASA distributions (Figure 3C,E). Such a discrepancy is explained by the initial increased intersheet spacing of class 3, dictated by the bigger volume of the enclosed aromatic rings (Figure 2). A comparable structural swelling is also observed for the PATH model. Associated R_g and SASA curves are perfectly coherent, with an initial decrease in the first 50 ns, followed by a steady increase for the rest of the simulation (Figure S3C,E), hence shifting its R_g and hydration shell toward higher values (Figure 3C,E). On the contrary, a decrease is observed for the class 1 topology within the first 100 ns on both graphs, unveiling a prompt gain in compactness that leads to a more tightly packed zipper.

The total intermolecular H bond content (Figure S5A) and number of backbone H bonds (Figure S5C) remain unsurprisingly very stable over the course of time. Remarkably, CORDAX and FibPred density profiles (Figure 4A,C) are nearly identical, showing that the interactions within the steric zipper remain vastly undisturbed. Along with STRIDE

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

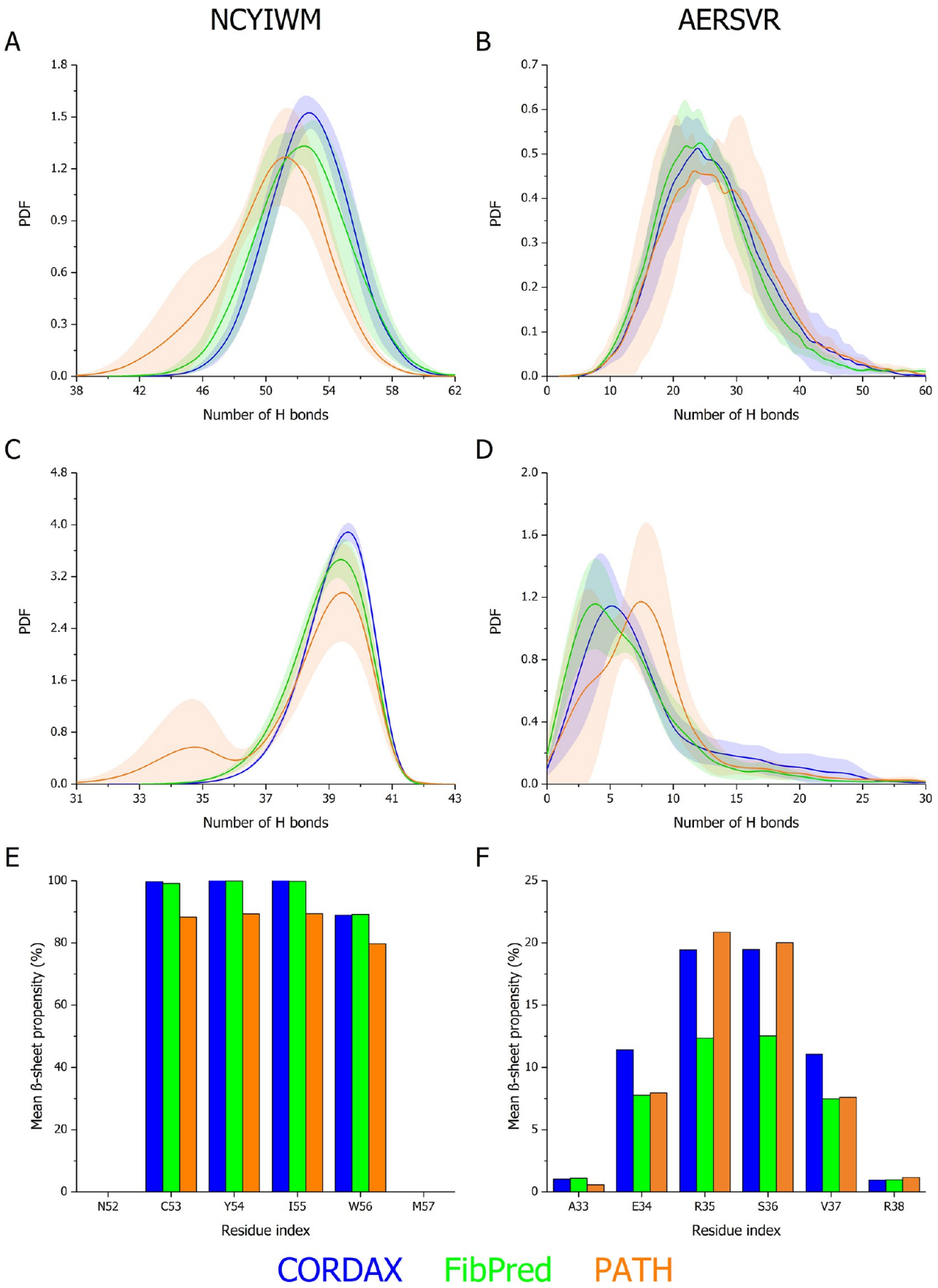


Figure 4. H bond network and secondary structure integrity of ^{52}NCYIWM^{57} and ^{33}AERSVR^{38} steric zippers. Probability density function (PDF) of (A, B) the total number of H bonds and (C, D) backbone H bonds, as well as (E, F) time-averaged per-residue β-sheet propensity of CORDAX (blue), FibPred (green), and PATH (orange) models of NCYIWM (A, C, E) and AERSVR (B, D, F) peptides simulated for 1 μs. On the first four

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

Figure 4. continued

panels (A−D), curves correspond to the average of triplicates with the standard deviation represented as a trace (shaded area) in the model-associated color. On the β-sheet propensity panels (E, F), per-residue values have been averaged over the 10 hexapeptide chains constituting the initial steric zipper.

assignment (Figures 4E and S5E), invariability in the proportion of backbone H bonds underscores the permanence of a β-sheeted structure, and the extensive network of interchain H bonds undergoes little to no disruption. Regarding the class 4 zipper, while its H bond content initially increases, it noticeably drops before plateauing at around 600 ns, a time step that corresponds to the RMSD rise attributed to a partial and local destructuration of one of the chains. The loss of a β-strand is consistent with the decrease observed in the β-sheet content of the PATH model by the end of the simulations (Figure S5E). Nevertheless, class 4 still maintains a high degree of H-bonded peptides most exclusively defined by β character. The latter motif is populated by C53, Y54, I55, and W56, expected to form the structural basis of the generated cross-β amyloids (Figure 4E). The previously pinpointed flexibility of the N-termini and associated ladder dynamics are notoriously exemplified by N52 presenting no β-sheet propensity whatsoever.

All of these trends become even more stringent by examining the changes that affect the structure of steric zippers after reducing their trajectory dimensionality (Figure 5). Strikingly, the central structure of each model populates, on average, 99.9% of the trajectory. Such a value is another unambiguous indicator of the ^{52}NCYIWM^{57} ability to form highly stable steric zippers irrespective to their specific topological features. For all of the classes, we can observe that the zippers mostly relaxed through intersheet crossing and interstrand twisting, which are accompanied by the realignment of several amino acids at the dry and wet interfaces (Figures 5, 6, and S6). For example, Asn (N52) side chain-backbone ladders are persistent and now captured on both the sheets. Along with the twisting, Tyr (Y54) side chains remarkably shift from a sandwich-like to a more favorable parallel displaced $\pi-\pi$ stacking, which is known to decrease their interaction energy, hence further stabilizing the amyloid zippers.[91]

Such transformations are responsible for the transient RMSD increase as well as the other discrepancies previously described. The higher R_g and SASA deviations observed for the FibPred model are mainly due to higher crossing angles between the sheets (Figure 5, middle left), as exemplified by the second population at around 51° (Figure 6A). Slight decompaction, hydration, and loss of H bonds for the PATH zipper arise from a higher degree of crossing (Figure 6A) and lability at the edges (e.g., unstructured chain I) combined with the one side opening of the zipper (Figure 5, bottom left). As seen from the top view and the radial distribution plots of water molecules with respect to the amino acids at the dry interface (Figure S7A), the FibPred core is more swollen and hydrated due to the needed intersheet distance arising from the size of enclosed aromatic rings.

The PATH zipper has also a larger SASA and slight water entry due to structure loosening and opening for accommodating more stable tyrosyl π stacking and reduce the steric hindrance of core aromatic side chains. Ile−Tyr CH−π interactions indeed convert into Tyr−Tyr parallel displaced $\pi-\pi$ stacking of lower energy. Interestingly, classes 3 and 4 result in zippers with a higher interstrand twist degree (Figure 6B), which may be induced by their more crowded dry interface and will have an impact on the macroscopic morphology of the resulting fibril. In comparison, CORDAX conserves a tighter, more interdigitated, and dryer Cys−Ile zipper arrangement (Figure 5, top left), excluding any water molecules from the core (Figure S7A). Overall decrease in R_g and SASA underlined at the beginning of the simulations advocates for such greater structural integrity and compactness. Moreover, the increase in the number of H bonds for the class 1 symmetry within the first 100 ns (Figure S5A,C) indeed corroborates with the Asn ladder extension as well as sporadic intrasheet Asn−Tyr N−H···O H bonds (Figure S6B).

Taken together, these results seemingly highlight a polymorphic susceptibility for the ^{52}NCYIWM^{57} stretch, which is relevant to the definition of its amyloidogenic character. Indeed, steric zipper and macroscopic polymorphism has extensively been captured by high-resolution structures for prototypical amyloidogenic proteins in their fibrillated state and is even recognized as participating in proteinopathies.[92,93] Furthermore, replicated simulations carried out on the ^{306}VQIVYK^{311} (PHF6) hexapeptide of the amyloidogenic protein tau serves as a positive control and further advocates for the pro-amyloid nature of the DPF3 ^{52}NCYIWM^{57} peptide, as they exhibit comparable zipper geometry and properties. Similarly, CORDAX predicts that ^{306}VQIVYK^{311} favors a class 1 steric zipper topology by burying Ile side chains and establishing a surface-exposed H-bonded amide ladder (here inter-Gln). This topology is also revealed to be highly stable over the course of time in terms of compaction, degree of hydration, and persistence of a β-sheeted network (Figures S10 and S11) and presents characteristics upon relaxation that are in good agreement with published experimental and simulation studies.[94] Indeed, the interstrand twist gain of ^{306}VQIVYK^{311} is superior to that of ^{52}NCYIWM^{57} which is consistent with the data collected from the high-resolution structures of PHF6 fibrils (Figure S12).

3.5. Simulation of Steric Zippers for a Nonamyloid Hexapeptide

Models constructed for the ^{33}AERSVR^{38} hexapeptide (Figure 2), consensually predicted as nonamyloid, were simulated over the course of replicated 1 μs trajectories. Whereas ^{52}NCYIWM^{57} zippers were characterized by an impressive overtime stability and integrity, those of ^{33}AERSVR^{38} tend to rapidly disorganize regardless of the starting topology.

Within the very first nanoseconds of the simulations, the same tendency is indeed observed among the different steric classes: a huge and steady increase in RMSD, R_g, and SASA (Figure S3B,D,F), accompanied by a significant, 10- to 20-fold, enhancement in per-residue (Figure 3B) and per-chain (Figure S4A−C) flexibility. This results in less core-discriminated residues along the sequence, as well as tremendously greater decompaction and hydration of all the chains (Figure 3D,F) compared to ^{52}NCYIWM^{57}. The data altogether point to the

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

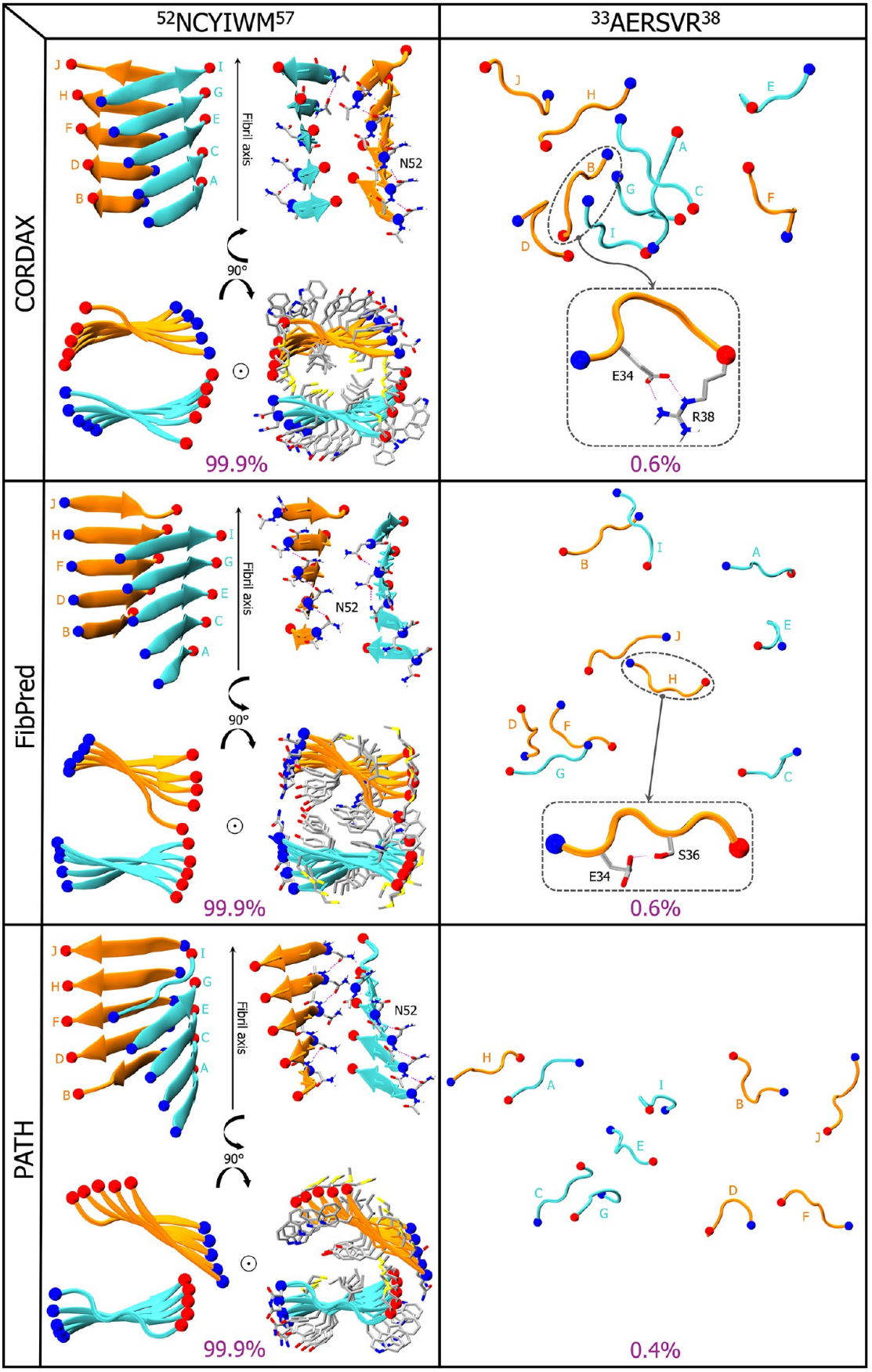


Figure 5. Representative structures of ^{52}NCYIWM^{57} and ^{33}AERSVR^{38} steric zippers after simulation. Presented snapshots correspond to the central structure of the most populated cluster among the triplicates, the population size percentage of which is indicated for the selected trajectory at the bottom of each panel. In the case of the AERSVR peptide, the final frame has been considered after clustering. On each panel, peptide chains

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

Figure 5. continued

(annotated from A to J) are displayed in cartoon representation with the N- and C-termini respectively pinpointed as blue and red spheres, the fibril axis is depicted by a black arrow, and the front and rear β-sheets initially assigned for the starting models are colored in cyan and orange, respectively. For every combination of the NCYIWM-modeler, the side view (top) is compared to the top view (bottom) of the resulting steric zipper. The side view, perpendicular to the fibril axis, shows the crossing between β-sheets (left) and the persistence of an Asn side chain-backbone ladder (magenta dashes, right), whereas the top view, down the fibril axis, evidences the β-sheet twisting (left) and the reorientation of amino acid side chains (right), highlighted in licorice representation.

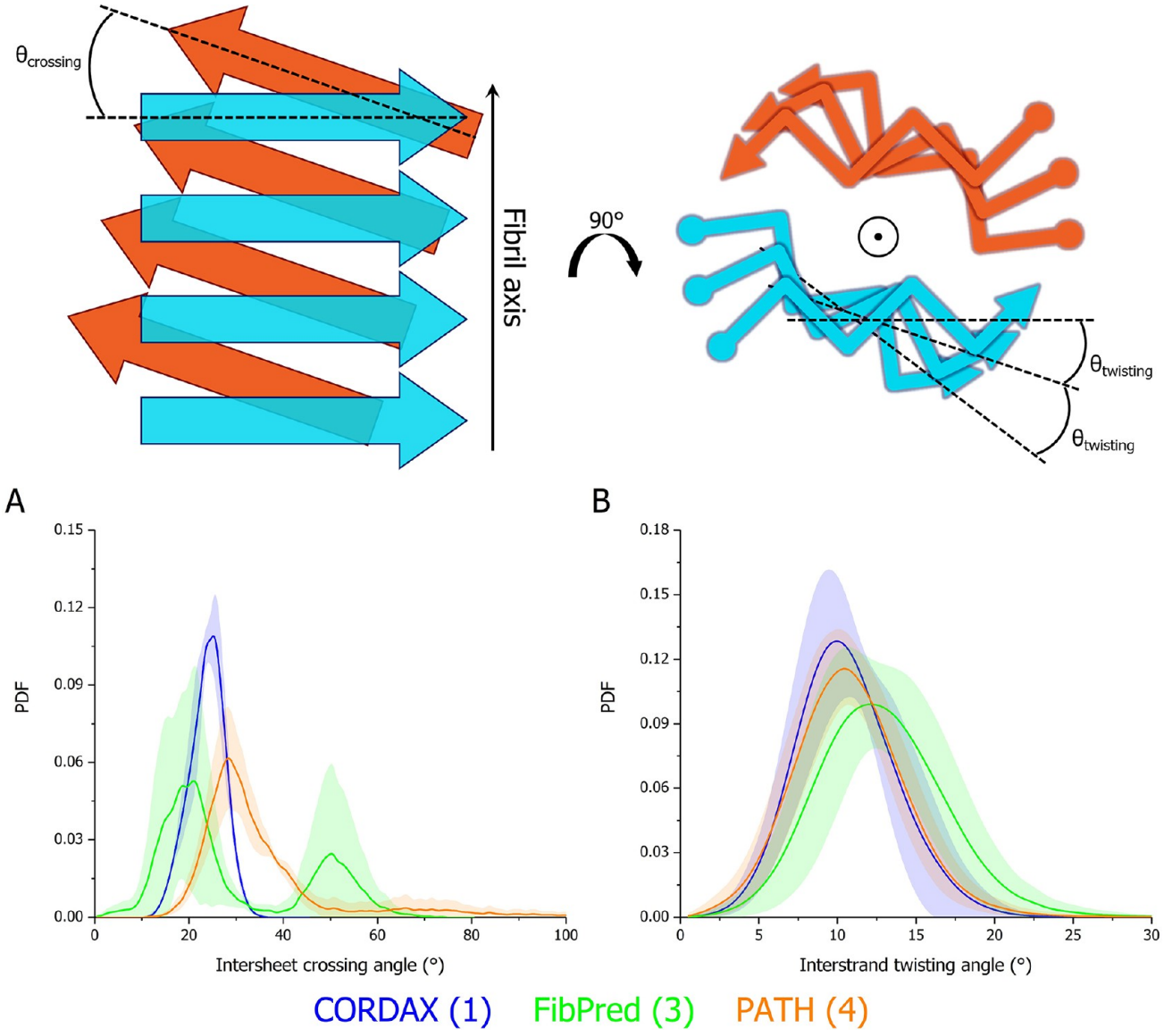


Figure 6. Crossing and twisting properties of ^{52}NCYIWM^{57} steric zippers. Schematic representation of the determined intersheet crossing angle (left) and interstrand twisting angle (right) in relaxed steric zippers (upper panel). Distribution of (A) intersheet crossing angle and (B) mean per-chain interstrand twisting angle (lower panel) is defined as a probability density function (PDF) of CORDAX (blue), FibPred (green), and PATH (orange) models simulated for 1 μs, with the corresponding steric zipper classes indicated in parentheses. Curves correspond to the average of triplicates with the standard deviation represented as a trace (shaded area) in the model-associated color.

fact that steric zippers have been dismantled and peptides are dispersed back free in solution.

Such claims are also substantiated by the evolution of the H bond number and secondary structures, which are, once more, very consistent. Intermolecular (Figure S5B) and backbone (Figure S5D) H bonds concomitantly decrease during the first 200 ns of the simulations, corresponding to the loss of β-sheet elements (Figure S5F). With respect to ^{52}NCYIWM^{57}, ^{33}AERSVR^{38} H bond populations are expectedly confined toward lower quantities (Figure 4B,D) and its composing amino acids display considerably reduced β-sheet occurrence (Figure 4F). Furthermore, radial distribution profiles are unequivocally informative about water entry due to zipper disassembly (Figure S7B). It can be pointed out that the computed parameters display low deviation and have very reproducible trends, showing that the loss of the initial steric

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

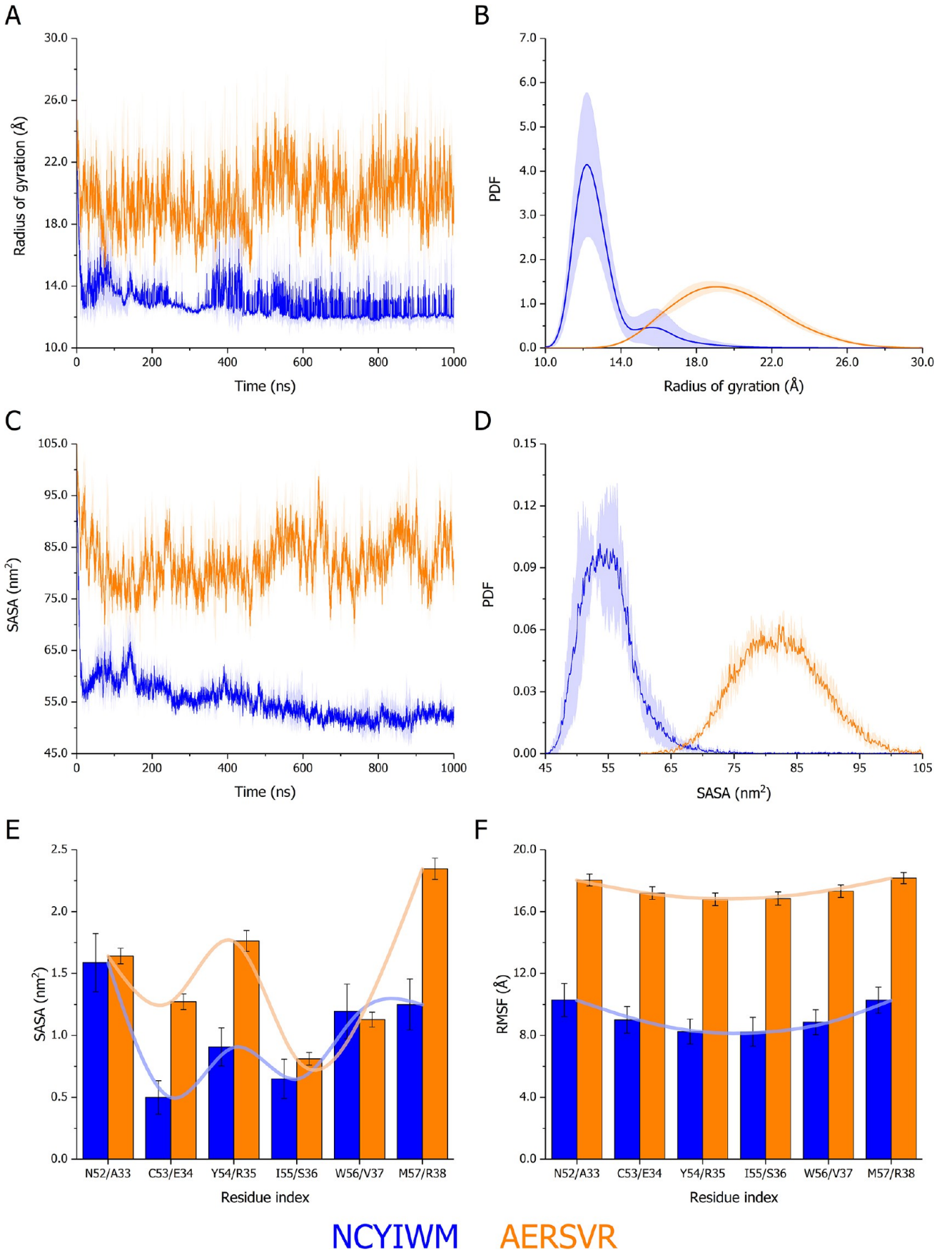


Figure 7. Compaction, hydration, and flexibility properties of ^{52}NCYIWM^{57} and ^{33}AERSVR^{38}free hexapeptide systems. Time evolution of (A) radius of gyration (R_g) and (C) SASA, as well as their respective distribution (B, D) defined as a probability density function (PDF), of NCYIWM (blue) and AERSVR (orange) free monomers (10 copies) simulated for 1 μs. Time-averaged per-residue (E) side chain SASA and (F) RMSF over

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

Figure 7. continued

the trajectories. On the R_g and SASA panels (A−D), curves correspond to the average of triplicates with the standard deviation represented as a trace (shaded area) in the peptide-associated color. On the SASA and RMSF panels (E, F), per-residue values have been averaged over the 10 dispersed hexapeptide chains, standard deviations between triplicates are depicted as error bars, and global trends along the peptide sequence as curved lines in the peptide-associated color.

zipper arrangement is a consistent phenomenon among the triplicates.

Visualization of the clustered trajectories further attest the unlikeness of ^{33}AERSVR^{38} to exist in highly organized cross-β architectures (Figure 5). While the most populated clusters are barely representative (∼0.4−0.6%), the final extracted frames reveal that the peptides adopt coil and turn structures involving only minimal backbone or side chain interactions between the chains. As seen on the CORDAX (Figure 5, top right) and FibPred (Figure 5, middle right) panels, ^{33}AERSVR^{38} monomers can populate bent conformations enriched in turn via intramolecular H bonds involving Glu with positively charged (Arg) or polar (Ser) amino acids. Relevantly, it has been reported that the fibrillation of Aβ can be impaired or even shut down by engineering monomers with more constrained and/or stronger H-bonded turns.[95] As such, ^{33}AERSVR^{38} proneness to adopt turn conformers can notably account for its relative inability to form extended β-sheets with its chains, the self-assembly of which is prevented by the dominating repulsion between the cationic guanidinium moieties of Arg residues (R35 and R38).

Notwithstanding the correlation existing between steric zipper stability and peptide amyloidogenicity, a direct link or causality cannot be drawn between those two properties. Indeed, several physicochemical, environmental, regulatory, and kinetic factors are at play for determining the (non)-amyloid profile of a given sequence. As steric zipper relaxation alone does not account for the nucleation driving step during the fibrillation process, the following simulations aim at remedying those limitations and further exploring the conformational space of the tested peptides.

3.6. Unbiased Self-Assembly Susceptibility of (Non)Amyloid Peptides

Although predicted models were found to be of high persistence during relaxation, modelers may introduce biased contacts that may not necessarily account for the peptide inherent amyloidogenicity, as they enforce a steric zipper configuration. To substantiate our previous results and assess to which extent the ^{52}NCYIWM^{57} segment is prone to aggregate and convert into β-sheets, we replicated the simulations consisting of 10 copies of the DPF3 hexapeptide in a random coil conformation and randomly dispersed in saline water. Taken as a nonamyloid referential for comparison purpose, a similar free hexapeptide system was constructed and simulated with the ^{33}AERSVR^{38} region.

Analysis of the trajectories uncovers that ^{52}NCYIWM^{57} hexapeptides cluster together under the first 10 ns, as evidenced by a significant reduction of the system R_g (Figure 7A) and SASA (Figure 7C). While small variations are observed along the simulations, attributed to the remanence of chain dynamics (Figure 7F), they eventually converge toward mean R_g (Figure 7B) and SASA (Figure 7D) values and distributions in the same order of magnitude as those of steric zippers. Such system compaction and chain desolvation are not observed for ^{33}AERSVR^{38}, presenting greater variations over time and broader populations toward higher R_g and SASA values. Furthermore, such discrepancies are reflected by per-residue and per-chain RMSF, showing that ^{33}AERSVR^{38} (Figures 7F and S8D) have a 2-fold increased flexibility compared to ^{52}NCYIWM^{57} peptides, the latter stiffening upon aggregation (Figures 7F and S8C). Across the ^{52}NCYIWM^{57} chains, Cys (C53) and Ile (I55) side chains exhibit lower solvent accessibility with an increased rigidity (Figure S8C) compared to aromatic residues and termini (Figures 7E and S8A), underlining a higher degree of burying. Intrachain RMSF trends between steric zippers (Figure S4) and free monomers (Figure S8) are globally in good agreement for the two hexapeptides.

Consistent across the replicates, the number of contacts (Figure S8E) and H bonds (Figure 8A) steadily increases over the course of time for ^{52}NCYIWM^{57}. After the initial clustering step presumably driven by hydrophobic packing of Cys and Ile amino acids, the peptides gradually stabilize an interchain H bond network. Whereas a similar increase is observed for ^{33}AERSVR^{38} during the first 100 ns, the curves exhibit greater variability with a significantly reduced number of contacts, resulting in a lower count of H bonds with a broader distribution (Figure 8B). Interestingly, most of the newly formed H-bonded pairs engage backbone atoms for ^{52}NCYIWM^{57} (Figure 8C,D), unveiling that the hexapeptides progressively undergo disorder-to-order transitions upon associating and interacting with each other. Strikingly, the time evolution profile of the number of backbone H bonds perfectly correlates with the creation of β-sheets, supporting the notion that the hit amyloid hexapeptide has a significant β-sheet proneness (Figure 8E).

Such a propensity is further ascertained by the time-averaged frequency of β character for each residue, disclosing a tendency similar to that of steric zippers with C53, Y54, I55, and W56 being the main components of the sheets (Figure 8F). Therefore, reduction in solvent accessibility and chain flexibility is associated with the structuring of ^{52}NCYIWM^{57} peptides into β-sheets. On the other hand, H bonds formed during the simulation of ^{33}AERSVR^{38} peptides are predominantly attributed to nonspecific interactions between charged and/or polar side chains without secondary structure generation. This is evidenced by the consequent shift in the PDF profile of backbone H bonds (Figure 8D) and the overall low per-residue β-sheet occupancy (Figure 8F). Although some β-sheets are sporadically formed, their content almost reaches 0% toward the end of the simulations (Figure 8E).

Trajectory clustering underpins such behavioral divergence between the two systems. By sorting the structures using a RMSD cutoff value of 5.0 Å, the number of resulting clusters straightforwardly converges for ^{52}NCYIWM^{57}, while ^{33}AERSVR^{38} fails to fall into populated basins due to high heterogeneity and chain dynamics (Figure S8F). This becomes even more conspicuous by extracting the structure corresponding to the most populated clusters, representing 99.6 and 0.3% of the trajectory for ^{52}NCYIWM^{57} and ^{33}AERSVR^{38},

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

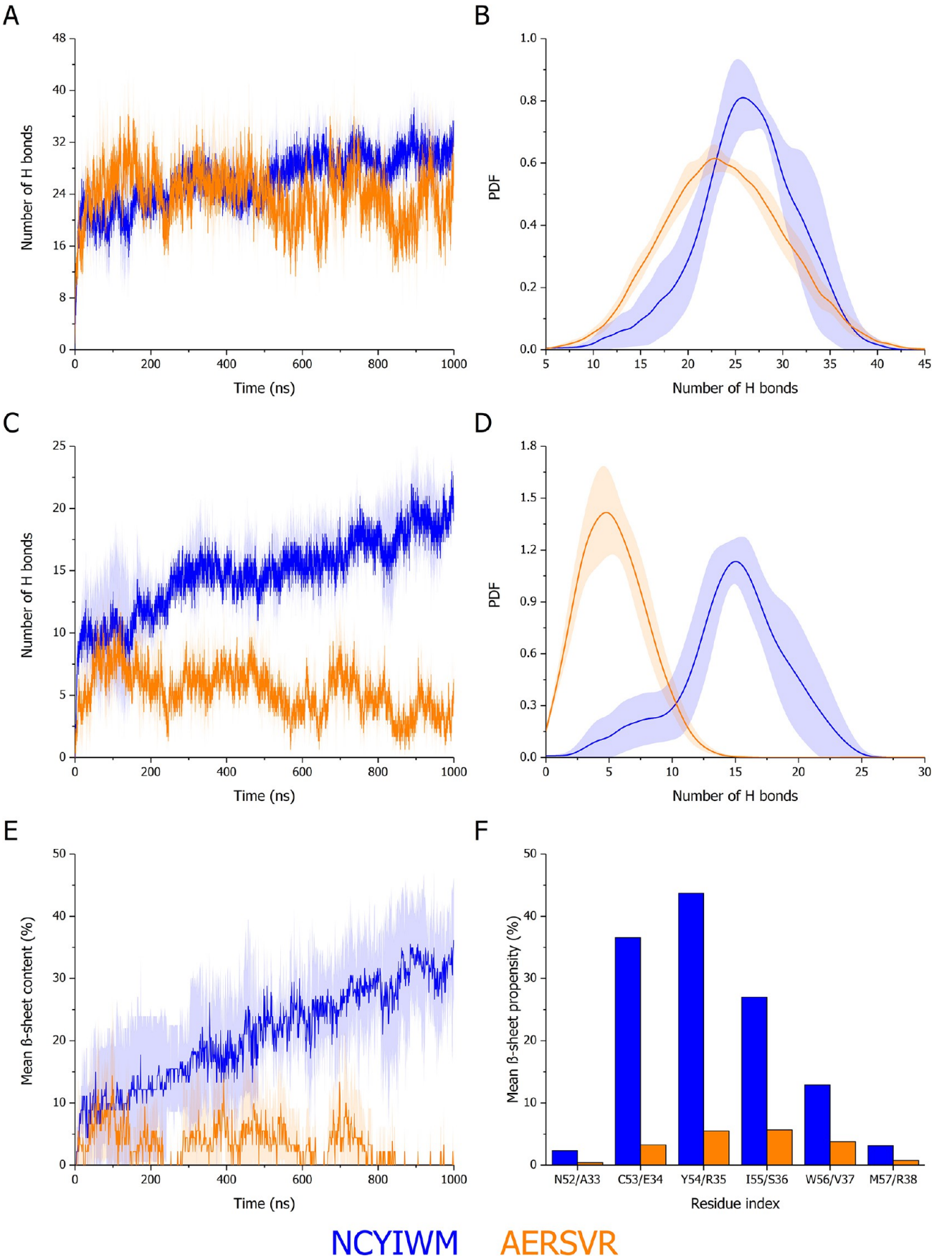


Figure 8. H bond network and secondary structure evolution of ^{52}NCYIWM^{57} and ^{33}AERSVR^{38}free hexapeptide systems. Time evolution of (A) the total number of H bonds and (C) number of backbone H bonds, as well as their respective distribution (B, D) defined as a probability density function (PDF), of NCYIWM (blue) and AERSVR (orange) free monomers (10 copies) simulated for 1 μs. (E) Time evolution of the β-sheet

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

Figure 8. continued

content and (F) time-averaged per-residue β-sheet propensity over the trajectories. On the first five panels (A−E), curves correspond to the average of triplicates with the standard deviation represented as a trace (shaded area) in the peptide-associated color. On the β-sheet propensity panel (F), per-residue values have been averaged over the 10 dispersed hexapeptide chains.

respectively (Figure 9A). In the latter case, notwithstanding partial compaction and ordering into β-bridges and turns, the hexapeptides primarily remain dispersed as statistical coil conformers.

Regarding ^{52}NCYIWM^{57}, the snapshot unambiguously shows that the peptides arrange into a β-sheeted oligomer composed of a dimer, a trimer, and a tetramer tightly packed together. Interestingly, within this heterogeneous oligomeric state, the β-strands preferentially adopt an antiparallel orientation (Figure 9A) compared with the parallel configuration predicted for the steric zippers (Figure 2). While such divergence may appear contradictory, it is, on the contrary, most likely informative on the amyloidogenic pathway of DPF3. Indeed, studies have underlined that, at the early amyloid stage, amyloidogenic polypeptides, such as Aβ, tend to aggregate mainly into antiparallel β-sheeted oligomers that later propagate into parallel fibrils during the elongation phase.[96−99] In that respect, the DPF3 ^{52}NCYIWM^{57} region is highly prone to nucleate into oligomers through the association of antiparallel β-sheet multimers. The latter first lump altogether before twining an interchain network of H bonds with additional chains, eventually elongating into parallel (proto)fibrils on a larger scale. Furthermore, inspection of the hydrophobic pocket formed at the oligomer center is coherent with the per-residue SASA analysis (Figure 7E) and suggests that rapid chain clustering is triggered by burying Ile and Cys side chains (Figure 9A, zoomed inset). In a similar fashion, formation of heterogeneous oligomers mediated by Ile residues has also been reported for the amyloidogenic PHF6 hexapeptide ^{306}VQIVYK^{311} derived from the tau protein, using all-atom MD simulations and starting from random coil conformers.[100] Regarding the PHF6 hexapeptide, we carried out additional simulations starting from free monomers of ^{306}VQIVYK^{311} in random coil conformation to compare its self-assembly behavior with DPF3 ^{52}NCYIWM^{57}. As for the modeled steric zipper, the results are once more convergent and strengthen not only our findings regarding ^{52}NCYIWM^{57} amyloidogenicity but also the overall applicability of our methodology to different protein systems. Similar tendencies are observed between the two peptides falling into the same order of magnitude, i.e., (i) fast clustering into a compact and dehydrated assembly (Figure S10), (ii) rapid H bond networking, and (iii) steady enrichment in β-sheets (Figure S11), all evolving toward the values obtained for the zipper precursor. Coherently with previous *in silico* reports, β-sheeted oligomers formed by the ^{306}VQIVYK^{311} peptide also present a predominant parallel orientation via aliphatic side chains (Ile and Val) burying (Figure S13).[101] On the contrary, the ^{52}NCYIWM^{57} hexapeptide favors an antiparallel orientation during nucleation (Figure 9A). Interestingly, this discrepancy underscores that, albeit sharing similar properties in terms of sequence composition and zipper geometry, tau and DPF3 exhibit distinct inherent nucleation behavior, which seemingly proves relevant to their fibrillation pathway.

To better appreciate the preferential states adopted by the two free hexapeptide systems, their conformational landscape was further explored by estimating free energy surfaces (FES), i.e., relative occupancies, along the two main coordinates obtained from the PCA of trajectories (Figure S9). ^{33}AERSVR^{38} exhibits a highly expanded FES centered around broad and connected basins in which peptide chains adopt coil and turn conformations in a highly hydrated state (Figure 9B, right). As previously stated, whereas the formation of β-sheets is sparsely detected for ^{33}AERSVR^{38}, they correspond to metastable dimers that eventually dissociate and more favorably stay free in solution or group into small fuzzy clusters. Conversely, free ^{52}NCYIWM^{57} stretches are highly unstable, and instead the peptides occupy well-separated and narrow wells of low energy, in which they are clumped and rearranged into tightly packed β-sheeted oligomers (Figure 9B, left). Interestingly, low free energy basins are associated with higher order multimers (i.e., addition of new β-strands), leading to the extension of β-sheets. This is likely due to the stabilizing H bond networking and side chain−side chain interactions.

From a morphological perspective, Asn, Tyr, and Trp residues have been shown to promote the formation of fibrous and elongated aggregates due to the directionality of H bonding and π stacking, while Cys, Ile, and Met amino acids induce more spherical morphologies. This is coherent with the hydrophobic effect favoring the reduction of the surface area of aggregated species in water solutions.[102] In the present case, compensating effects are likely at play, each one dominating at certain stages of fibril growth. The mixed composition of ^{52}NCYIWM^{57} leads to round-shaped aggregates, which nonetheless have chains aligning to some extent into fibrous structures similar to twisted β-barrels, such as those reported for Aβ16-22-PHF6 hetero-oligomers with contacts between hydrophobic side chains, including Ile.[20] In light of such preferences and the results discussed so far, we propose that nucleation of the amyloid hexapeptide ^{52}NCYIWM^{57} is initially driven by the hydrophobic packing of Ile and Cys side chains. This first step leads to the generation of pseudospherical antiparallel β-sheet oligomers that further transition to parallel β-sheet zipper structures upon elongating into larger and fiber-like supramolecular assemblies. At this stage, the directionality and stability effect of Asn ladders and Tyr π stacking become prevalent.

3.7. Seeding Ability of a Relaxed Precursor for Monomer Nucleation and Fibril Elongation

While the hit amyloid hexapeptide was shown to stabilize steric zippers of diverse topologies and to inherently possess the capacity to rearrange into β-sheet oligomers, we also examined if the steric zipper could serve as a template for inducing the conversion of free monomers into the fibrillar structure. In that purpose, the previously relaxed steric zipper structure with the highest overtime integrity, that is, the CORDAX class 1 model, was placed at the center of a saline water box and surrounded by 10 copies of the ^{52}NCYIWM^{57} hexapeptide in random coil conformation and at random positions.

The overall system properties (Figure 10) are very convergent as well as consistent with the simulations

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

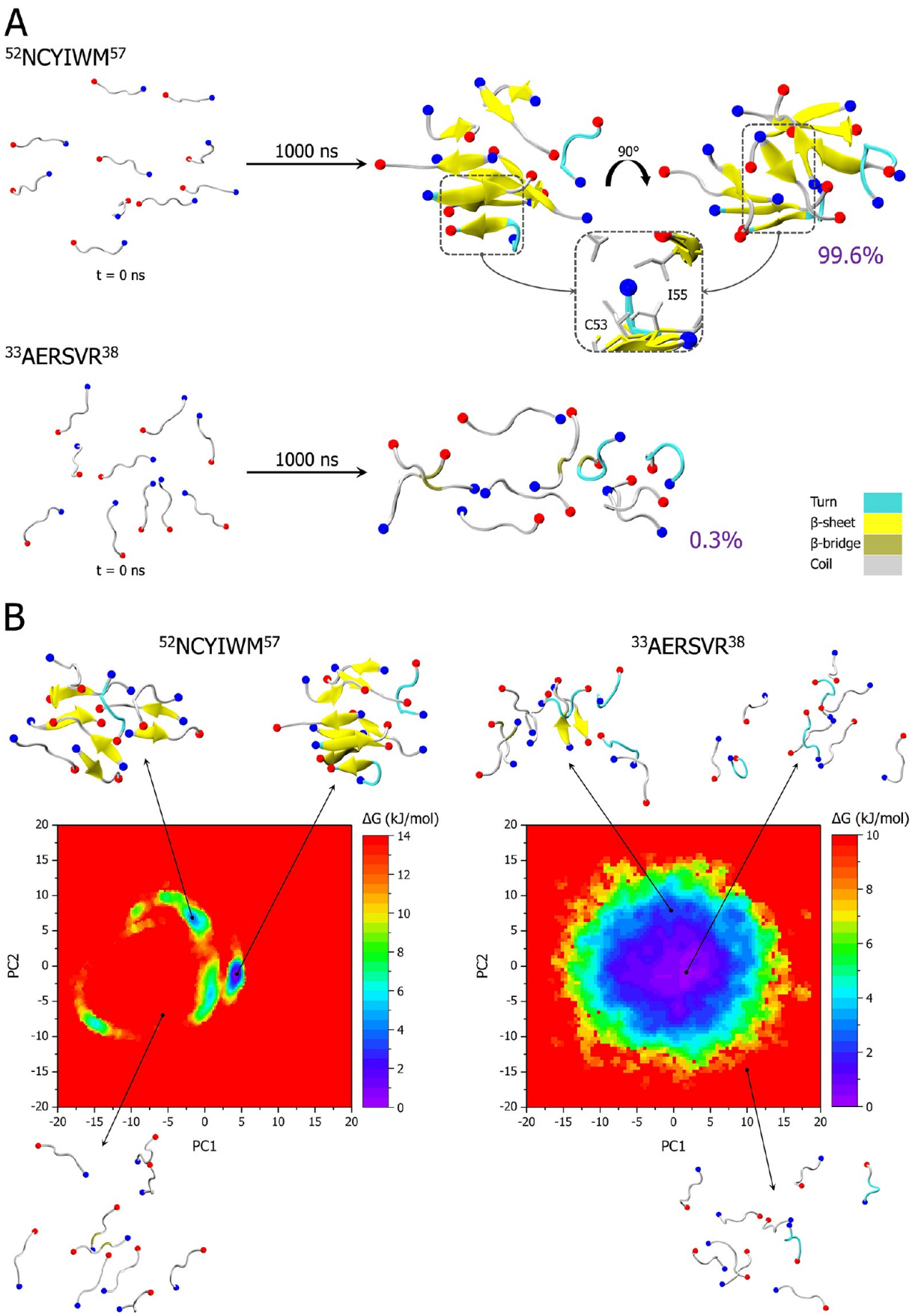


Figure 9. Structural transition and conformational landscape of ^{52}NCYIWM^{57} and ^{33}AERSVR^{38}free hexapeptide systems. (A) Representative snapshots of NCYIWM (top) and AERSVR (bottom) free monomers (10 copies) at the start of the simulation (left) and after 1 μs (right). Final structures correspond to the central structure of the most populated cluster among the triplicates, the population size percentage of which is

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

Figure 9. continued

indicated in purple for each selected trajectory. (B) Free energy surface (FES) of NCYIWM (left) and AERSVR (right) systems projected along the first two principal components (PC1 and PC2) of the selected trajectory. On each panel, peptide chains are shown in cartoon representation with the N- and C-termini respectively pinpointed as blue and red spheres, and turn, β-sheet, β-bridge, and coil secondary structures colored in teal, yellow, gold, and gray, respectively.

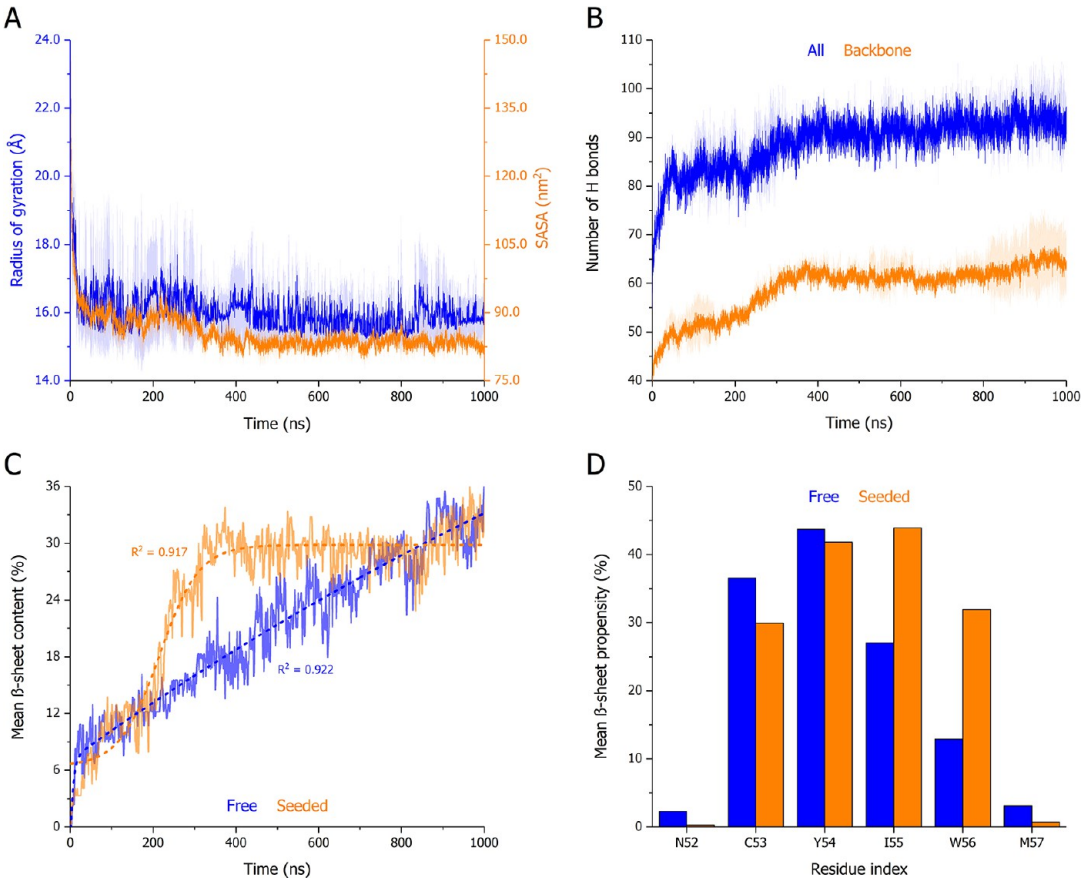


Figure 10. Compaction, hydration, and secondary structure evolution of the ^{52}NCYIWM^{57} seeded hexapeptide system. Time evolution of (A) the radius of gyration (R_g) coupled with the SASA and (B) total number of backbone H bonds of the whole system. (C) Time evolution of the β-sheet content and (D) time-averaged per-residue β-sheet propensity of the seeded free monomers (10 copies) over the trajectories. On the first two panels (A, B), curves correspond to the average of triplicates with the standard deviation represented as a trace (shaded area). On the β-sheet content panel (C), kinetic traces in the absence and presence of a precursor are fitted to exponential and logistical functions, respectively; standard deviation are not represented for better readability. On the β-sheet propensity panel (D), per-residue values have been averaged over the 10 dispersed hexapeptide chains.

comprising only free monomers (Figures 7 and 8). Indeed, peptides rapidly cluster together into compact and hydrophobic assemblies (Figure 10A) that are enriched in intermolecular contacts (Figure 10B), particularly through the formation of interchain β-sheets (Figure 10C). While the tendencies appear well reproduced with the precursor-free simulations and therefore could strictly arise from the spontaneous generation of isolated oligomers, interaction-associated and secondary structure curves plateau from 300 ns onward (Figure 10B,C). The content in the β-sheet for the added hexapeptides also increases at a faster rate (Figure 10C) to reach a similar percentage (around 30−35%) compared to the free monomer simulations (Figure 8E). Without seeding, β-sheet growth occurs exponentially within the first 25 ns before following a steady linear-like regime. Upon seed addition, the growth profile becomes sigmoidal, which can be described by an autocatalytic model. Moreover, the monomers show a more balanced β-sheet occupancy for their central amino acids (Figure 10D), with a noticeable increased propensity for I55 and W56, as well as a decrease for the residues at the edges, thus better echoing the profile obtained for the steric zippers (Figure 4E). Such observations strongly suggest that the amyloid zipper accelerates and enhances, hence catalyzes, the conversion of monomers into β-sheets through the precursor growth.

Examination of the structural states obtained throughout a selected trajectory at distinct time steps confirmed the seeding hypothesis (Figure 11). In the very first nanoseconds, two chains (L and T) assemble into a parallel β-sheet dimer in the vicinity of the steric zipper (step 1). In the following 40 ns, the dimer becomes a parallel trimer (addition of chain Q), with only the two initial chains being in-register, and approaches one of the edges of the zipper in an upward open orientation with a 63° angle (step 2). Under the same period, two additional chains (M and R) start to fold into an out-of-register antiparallel β-sheet dimer, and all of the chains are found in

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

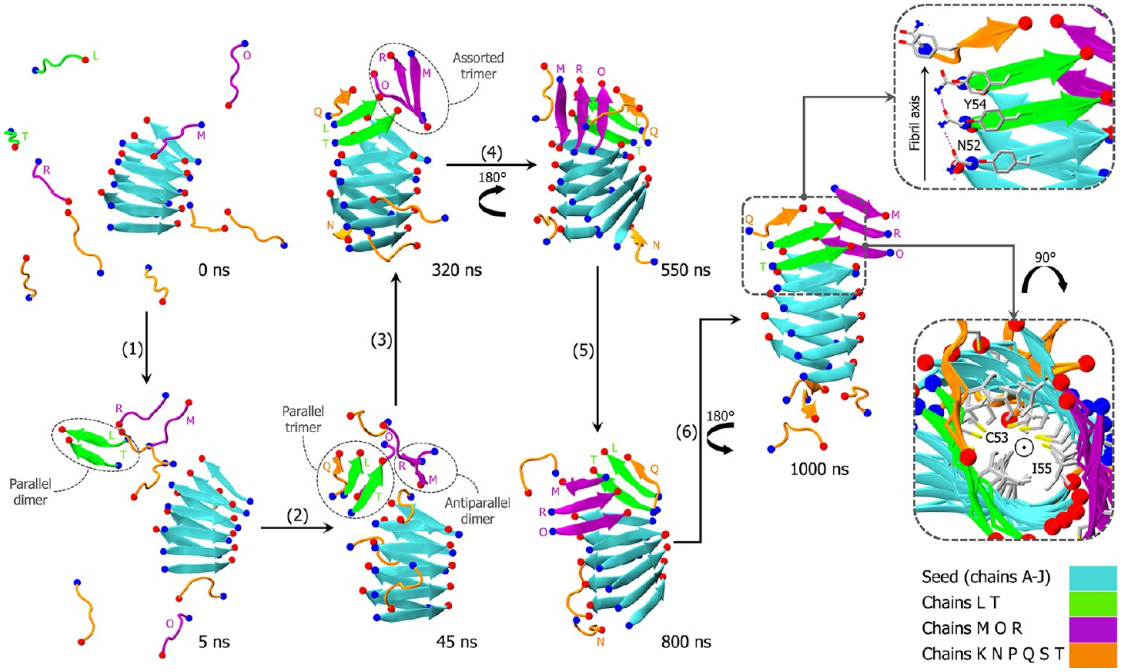


Figure 11. Fibrillation and elongation mechanism of the ^{52}NCYIWM^{57} hexapeptide. From a representative trajectory, selected snapshots reveal the structural rearrangements that occur at the level of the seeded monomers, highlighting the formation of distinct β-sheeted oligomer species (steps 1−3), able to interact with the relaxed steric zipper (steps 3−5) and leading to the elongation of the fibrillar architecture (step 6) while conserving a class 1 geometry through the Asn ladder extension and Tyr parallel displaced $\pi-\pi$ stacking (zoomed insets). Peptide chains are shown in cartoon representation with the N- and C-termini respectively pinpointed as blue and red spheres. In the zoomed insets, annotated amino acid side chains are displayed in licorice representation and H bonds as magenta dashes.

proximity with the zipper. Subsequently, the parallel trimer aligns in-register with one of the zipper sheets through the creation of new interchain backbone H bonds, while the antiparallel dimer extends to a mixed trimer by adding up one parallel strand (chain O) at its extremity oriented toward the precursor (step 3). From 320 ns onward, a similar two-step addition mechanism is observed for the assorted trimer: the latter approaches in an upward open orientation, forming a 65° crossing angle with the targeted β-sheet (step 4), before flipping down toward and anchoring to the said sheet via the trimer side consisting of in-register parallel β-strands (step 5). After 1 μs, the final snapshot remarkably shows that, out of the 10 monomers, 5 are found elongating the steric zipper edge via parallel β-sheets (step 6), an occurrence that was not observed in the precursor-free simulations.

Noteworthily, the orientation of Tyr (Y54) and Asn (N52) side chains reveals that the class 1 symmetry is maintained upon the addition of trimers to the steric zipper, and such contacts are likely involved in the templating process. Indeed, the multimer rotation and alignment come with the extension of the Asn N−H···O ladder and parallel displaced $\pi-\pi$ stacking of Tyr aromatic rings as the zipper is interestingly elongated by conserving its interstrand twisting feature. While out-of-register, one of the partially rearranged peptides (chain Q) of the parallel trimer already has its tyrosyl moiety in a more stackable orientation. In addition, the more favorable hydrophobic packing at the dry interface of Cys (C53) and Ile (I55) side chains observed in precursor-free and steric zipper simulations is also respected here for the added trimers thanks to the conservation of a class 1 topology. In that respect, we observe a comparable pathway reported for the oligomerization of amyloid DFNKF pentapeptides, derived from the human calcitonin, in which Asn ladder formation and extension are an important driving force of fibril elongation.[103] Asn residues act as a glue to stick the peptides with one another and direct their parallel assembly, and it was shown that upon asparagine-to-alanine mutation, the peptides lost their β-sheet propensity.

Remarkably, the antiparallel-to-parallel switch occurring only at larger and later fibrillation scales and stages has been substantially captured by our seeded simulations. The introduction of a parallel template (i.e., the relaxed class 1 steric zipper) in the simulation box shifted the preferential antiparallel oligomerization of ^{52}NCYIWM^{57} (Figure 9) toward the generation of parallel multimers that anchored to the zipper edges. As evidenced on several frames (steps 3 and 4), the direct addition of free monomers parallel to the precursor edge (chain N) cannot be excluded, but the process appears less competitive, and the added monomer sporadically disaggregates from the zipper. From a global perspective, our results indicate that the selected steric zipper acts as a template or seed that triggers the conversion of β-sheet multimers for fibril elongation by inducing a class 1 and parallel chain reorientation.

4. CONCLUSIONS AND PERSPECTIVES

We aimed at presenting an aggregation-oriented computational methodology based on the combination of freely available sequence-based predictors and modeling tools with all-atom classical MD simulations for the identification of pro-amyloid segments in proteins, for which structures are currently not solved or vexatiously elude *in vitro* characterization. Through selected examples, our multilevel approach promisingly showed the capability to precisely identify APRs and amyloid peptides, as well as to decipher the molecular mechanisms of peptide self-assembly at the atomistic scale.

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

By submitting the sequence of our protein of interest, i.e., the intrinsically disordered human epigenetic regulator DPF3a, to the pipeline, one hexapeptide, namely, ^{52}NCYIWM^{57}, stood out as one of the most pro-amyloid stretches. While all the modelers assumed the peptide to adopt antifacial parallel steric zippers stabilized through similar effects, they assigned distinct class symmetries according to the orientation of β-sheets and strands. Over 1 μs-simulation triplicates, we consistently observed that the ^{52}NCYIWM^{57} peptide retained a stable, rigid, and dry amyloid core. The integrity of different steric zipper classes likely underlines the susceptibility of the hexapeptide to aggregate into fibrillar polymorphs, which is another hallmark of amyloid aggregates. Supporting such a claim, our simulations pinpointed discrepancies in intersheet crossing and interstrand twisting angles among the relaxed zippers, which would result in distinct macroscopic properties, such as variations in fibril width and crossover distance that could be verified *in vitro*. As such, we advise exploiting multiple modelers (CORDAX, FibPred, and PATH) that structurally complement each other for assessing polymorphism and giving more weight to amyloid predictions.

Driven by the hydrophobic packing of Ile and Cys side chains, the ^{52}NCYIWM^{57} peptide has been shown to spontaneously and rapidly self-associate into antiparallel β-sheeted oligomers upon starting from free monomers in random coil conformation. After introduction of the relaxed class 1 steric zipper in the free monomer box, the amyloid precursor remarkably served as a templating structure that induced and accelerated the rearrangement of oligomeric species. Indeed, seeding reshaped the structural preferences of ^{52}NCYIWM^{57} by favoring parallel β-sheets. By combining all the simulation data, we were able to decipher the self-assembly mechanism of the amyloid hexapeptide ^{52}NCYIWM^{57} and propose the following fibrillation pathway. During the lag phase, free peptide monomers preferentially nucleate into antiparallel β-sheet oligomers, the subsequent association of which triggers an antiparallel-to-parallel switch, initiating the formation of an amyloid precursor arranged into a parallel and dry steric zipper. During the growth phase, the steric zipper shifts the conformational ensemble of free peptides toward parallel multimers that anchor to the precursor edge in an upward open orientation before flipping down. Multimer addition eventually elongates the amyloid structure by conserving the initial class symmetry through side chains (re)alignment. Behavioral convergence of ^{52}NCYIWM^{57} with the ^{306}VQIVYK^{311} (PHF6) amyloid peptide of tau not only advocates for the former amyloidogenicity but also strengthens the applicability and transferability of our protocol.

Therefore, compelling evidence designates the ^{52}NCYIWM^{57} hexapeptide found in the two isoforms of the DPF3 protein as an APR that exhibits a particularly high propensity to aggregate into cross-β amyloid structures. Thanks to its unique composition, nucleation and fibril elongation are enabled through key stabilizing and seeding interactions, such as hydrophobic packing of aliphatic or sulfydryl side chains, H-bonded Asn ladder, and $\pi-\pi$ stacking of the aromatic rings. Comparatively, stretches consensually predicted as nonamyloid, such as ^{33}AERSVR^{38}, are incapable of sustaining highly organized steric zippers or self-assembling into β-rich oligomers or multimers due to interchain electrostatic repulsion and steric hindrance, maintaining the peptides in a highly hydrated and statistical coil state. Taken together, these findings invite investigating such properties experimentally for the two hexapeptides, as well as assessing small molecules that could preclude DPF3 aggregation or destabilize its amyloid precursors in neurodegenerative contexts. Indeed, biophysical aggregation studies will be needed to fully ascertain if the ^{52}NCYIWM^{57} hexapeptide is necessary and/or sufficient for DPF3 isoforms fibrillation.

We acknowledge that solving fibrillar structures proceeds to unveil amyloid interfaces of increased complexity and variety, notably consisting of longer chains and/or different proteins involved in cross-interactions. In the purpose of simulating such scaled up and heterotypic interfaces, our methodology could integrate coarse grained (CG) simulation methods, as well as other specifically designed modeling tools, such as PACT and CABS-DOCK, that recently succeeded in predicting the structure of protofibrils.[104−106] By reducing the number of modeled atoms, CG could also help to set up simulation systems in more physiologically relevant conditions regarding, for example, protein concentration and cellular crowders. In addition, CG used in combination with time-lagged independent component analysis (TICA) would expand our understanding of the peptide aggregation mechanism through the reduction of dynamic system dimensionality, allowing the construction of Markov state models.[107] Moreover, noncanonical cross-α fibrils emerge as a new paradigm in the field as well as an important class of functional amyloids. As exploited modelers exclusively rely on extensive cross-β training data sets, our approach finds its own limitation and is untransferable to cross-α aggregates, which will have to wait for the development of innovative cross-α-oriented predictors, even though Cross-Alpha-Det and AlphaFold2 (ColabFold) hold promise in this regard.[108,109]

In a nutshell, our exclusively *in silico* approach proved to be reliable not only to detect APRs in proteins but also to discriminate polymorphic amyloid regions from nonamyloid peptides while opening alternative ways for elucidating the fibrillation mechanisms of complex protein systems that challenge experimental assays. We are hoping that such a strategy will help in the evaluation of new anti-amyloid agents targeting APRs in neurodegeneration or amyloidosis, as well as in the design of engineered peptides for the generation of novel amyloid-inspired biomaterials.

■ ASSOCIATED CONTENT

Data Availability Statement

All software, algorithms, and codes exploited in this manuscript are freely available online. Along with the Supporting Information, data, files, and in-house scripts have been deposited on and are accessible from the Zenodo repository under the following record ID: 19472089 (DOI: 10.5281/zenodo.19472089).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.6c00951.

The summary of the simulation parameters and metrics for the NCYIWM, AERSVR, and VQIVYK peptide systems; sequence-based predictions of APRs and amyloid peptides in DPF3a; structural highlights of NCYIWM and AERSVR steric zipper models before and after simulation; time evolution of RMSD, R_g, SASA, number of H bonds, and β-sheet content of NCYIWM and AERSVR steric zippers; time-averaged RMSF of NCYIWM and AERSVR steric zippers; radial distribu-

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

tion function g(r) of water molecules around NCYIWM and AERSVR steric zipper cores; time-averaged SASA and RMSF, time evolution of the number of contacts and clusters, and PCA profiles of NCYIWM and AERSVR free hexapeptides; distribution and time evolution of R_g, SASA, number of H bonds, and β-sheet content of the VQIVYK steric zipper and free hexapeptides; and structural highlights of the VQIVYK steric zipper and free hexapeptides before and after simulation (PDF)

AUTHOR INFORMATION

Corresponding Author

Julien Mignon − *Laboratoire de Chimie Physique des Biomolécules, UCPTS, University of Namur, Namur 5000, Belgium; Namur Institute of Structured Matter (NISM), Namur Research Institute for Life Sciences (NARILIS), University of Namur, Namur 5000, Belgium;* orcid.org/0000-0003-1223-4501; Phone: +3281724523; Email: julien.mignon@unamur.be

Authors

Tanguy Leyder − *Laboratoire de Chimie Physique des Biomolécules, UCPTS, University of Namur, Namur 5000, Belgium; Namur Institute of Structured Matter (NISM), Namur Research Institute for Life Sciences (NARILIS), University of Namur, Namur 5000, Belgium*

Hugo Bâlon − *Laboratoire de Chimie Physique des Biomolécules, UCPTS, University of Namur, Namur 5000, Belgium; Namur Institute of Structured Matter (NISM), Namur Research Institute for Life Sciences (NARILIS), University of Namur, Namur 5000, Belgium;* orcid.org/0000-0003-4315-0562

Antonio Monari − *Université Paris Cité and CNRS, ITODYS, Paris 75006, France;* orcid.org/0000-0001-9464-1463

Denis Mottet − *Molecular Analysis of Gene Expression (MAGE) Laboratory, GIGA Institute, University of Liège, Liège 4000, Belgium*

Catherine Michaux − *Laboratoire de Chimie Physique des Biomolécules, UCPTS, University of Namur, Namur 5000, Belgium; Namur Institute of Structured Matter (NISM), Namur Research Institute for Life Sciences (NARILIS), University of Namur, Namur 5000, Belgium;* orcid.org/0000-0001-8208-3310

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.6c00951

Author Contributions

J.M. conceptualized and designed the computational methodology and pipeline, set up the appropriate environments, conducted all the predictions, modeling, and simulations, and processed and analyzed the data, as well as wrote the original draft of the manuscript. C.M. supervised the work in the capacity of the laboratory director. A.M. and C.M. validated the processed data, as well as its interpretation. T.L., H.B., A.M., D.M., and C.M. reviewed and edited the manuscript. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful to the PTCI high-performance computing resource of the University of Namur (UNamur). The present research benefited from computational resources provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Belgian National Fund for Scientific Research (F.R.S.-FNRS) under grant no. 2.5020.11 and by the Walloon Region, and made available on Lucia, the Tier-1 supercomputer of the Walloon Region, infrastructure funded by the Walloon Region under the grant agreement no. 1910247. J.M. thanks the FNRS for his Research fellow fellowship. T.L. and H.B. also thank the FNRS for their Fund for Research training in Industry and Agriculture (FRIA) Doctoral grant. A.M. thanks ANR and CGI for their financial support of the present work through Labex SEAM ANR 11 LABEX 086 and ANR 11 IDEX 05 02. The support of the IdEx "Université Paris 2019" ANR-18-IDEX-0001 is also acknowledged. D.M. and C.M. are thankful to the FNRS for their Senior Research Associate position. J.M. is especially appreciative of the Switch Laboratory from the Katholieke Universiteit Leuven (KU Leuven, Belgium) for sharing the open-source code and standalone package of CORDAX, as well as of Jakub W. Wojciechowski from the Wroclaw University of Science and Technology (Poland) for his help in setting up PATH and sharing his insights.

ABBREVIATIONS

APR, aggregation prone region; Aβ, amyloid β; CG, coarse grain; DPF3, double PHD fingers 3; FibPred, FibPredictor; FES, free energy surface; MD, molecular dynamics; PME, particle mesh Ewald; PBC, periodic boundary conditions; PC, principal component; PCA, principal component analysis; PDF, probability density function; R_g, radius of gyration; RMSD, root-mean-square deviation; RMSF, root-mean-square fluctuation; SASA, solvent accessible surface area.

REFERENCES

(1) Bayer, T. A. Proteinopathies, a Core Concept for Understanding and Ultimately Treating Degenerative Disorders? *Eur. Neuropsychopharmacol.* **2015**, *25* (5), 713−724.

(2) Ajmal, M. R. Protein Misfolding and Aggregation in Proteinopathies: Causes, Mechanism and Cellular Response. *Diseases* **2023**, *11* (1), 30.

(3) Abubakar, M. B.; Sanusi, K. O.; Ugusman, A.; Mohamed, W.; Kamal, H.; Ibrahim, N. H.; Khoo, C. S.; Kumar, J. Alzheimer's Disease: An Update and Insights Into Pathophysiology. *Front. Aging Neurosci.* **2022**, *14*, 742408.

(4) Vidović, M.; Rikalovic, M. G. Alpha-Synuclein Aggregation Pathway in Parkinson's Disease: Current Status and Novel Therapeutic Approaches. *Cells* **2022**, *11*, 1732.

(5) Chatani, E.; Yuzu, K.; Ohhashi, Y.; Goto, Y. Current Understanding of the Structure, Stability and Dynamic Properties of Amyloid Fibrils. *Int. J. Mol. Sci.* **2021**, *22* (9), 4349.

(6) Muchtar, E.; Dispenzieri, A.; Magen, H.; Grogan, M.; Mauermann, M.; McPhail, E. D.; Kurtin, P. J.; Leung, N.; Buadi, F. K.; Dingli, D.; Kumar, S. K.; Gertz, M. A. Systemic Amyloidosis from A (AA) to T (ATTR): A Review. *J. Int. Med.* **2021**, *289* (3), 268−292.

(7) Sønderby, T.; Najarzadeh, Z.; Otzen, D. Functional Bacterial Amyloids: Understanding Fibrillation, Regulating Biofilm Fibril Formation and Organizing Surface Assemblies. *Molecules* **2022**, *27* (13), 4080.

(8) Garcia-Sherman, M. C.; Hamid, S. A.; Jackson, D. N.; Thomas, J.; Lipke, P. N. Functional Amyloids in Adhesion of Non-Albicans Candida Species. *Pathogens* **2025**, *14* (8), 723.

(9) Calabrese, A. N.; Liu, Y.; Wang, T.; Musgrave, I. F.; Pukala, T. L.; Tabor, R. F.; Martin, L. L.; Carver, J. A.; Bowie, J. H. The Amyloid Fibril-Forming Properties of the Amphibian Antimicrobial Peptide Uperin 3.5. *ChemBioChem* **2016**, *17* (3), 239−246.
(10) Bücker, R.; Seuring, C.; Cazey, C.; Veith, K.; García-Alai, M.; Grünewald, K.; Landau, M. The Cryo-EM Structures of Two Amphibian Antimicrobial Cross-β Amyloid Fibrils. *Nat. Commun.* **2022**, *13*, 4356.
(11) Qi, X.; Wang, Y.; Yu, H.; Liu, R.; Leppert, A.; Zheng, Z.; Zhong, X.; Jin, Z.; Wang, H.; Li, X.; Wang, X.; Landreh, M.; A. Morozova-Roche, L.; Johansson, J.; Xiong, S.; Iashchishyn, I.; Chen, G. Spider Silk Protein Forms Amyloid-Like Nanofibrils through a Non-Nucleation-Dependent Polymerization Mechanism. *Small* **2023**, *19* (46), 2304031.
(12) Bonar, L.; Cohen, A. S.; Skinner, M. M. Characterization of the Amyloid Fibril as a Cross-β Protein. *Proc. Soc. Exp. Biol. Med.* **1969**, *131* (4), 1373−1375.
(13) Almeida, Z. L.; Brito, R. M. M. Structure and Aggregation Mechanisms in Amyloids. *Molecules* **2020**, *25* (5), 1195.
(14) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A.; Riekel, C.; Eisenberg, D. Atomic Structures of Amyloid Cross-β Spines Reveal Varied Steric Zippers. *Nature* **2007**, *447*, 453−457.
(15) Sulyok-Eiler, M.; Harmat, V.; Perczel, A. Unravelling the Complexity of Amyloid Peptide Core Interfaces. *J. Chem. Inf. Model.* **2024**, *64*, 8628−8640.
(16) Stroud, J. C. The Zipper Groups of the Amyloid State of Proteins. *Acta Crystallogr. D Biol. Crystallogr.* **2013**, *69* (4), 540−545.
(17) Yang, J.; Agnihotri, M. V.; Huseby, C. J.; Kuret, J.; Singer, S. J. A Theoretical Study of Polymorphism in VQIVYK Fibrils. *Biophys. J.* **2021**, *120* (8), 1396−1416.
(18) Do, T. D.; Sangwan, S.; de Almeida, N. E. C.; Ilitchev, A. I.; Giammona, M.; Sawaya, M. R.; Buratto, S. K.; Eisenberg, D. S.; Bowers, M. T. Distal Amyloid β-Protein Fragments Template Amyloid Assembly. *Protein Sci.* **2018**, *27* (7), 1181−1190.
(19) Agha, M. M.; Uversky, V. N. Chapter Four—Morphological Features and Types of Aggregated Structures. In *The Hidden World of Protein Aggregation*; Dabirmanesh, B., Teplow, D. B., Uversky, V. N. ., Eds.; Academic Press, 2024; Vol. *206*, pp 85−109.
(20) Li, X.; Chen, Y.; Yang, Z.; Zhang, S.; Wei, G.; Zhang, L. Structural Insights into the Co-Aggregation of Aβ and Tau Amyloid Core Peptides: Revealing Potential Pathological Heterooligomers by Simulations. *Int. J. Biol. Macromol.* **2024**, *254*, 127841.
(21) Dey, S.; Kumar, R.; Mishra, R.; Bera, S. Exploring Cross-α Amyloids: From Functional Roles to Design Innovations. *Trends Biochem. Sci.* **2024**, *49* (12), 1097−1110.
(22) Nowick, J. S.; Insaf, S. The Propensities of Amino Acids To Form Parallel β-Sheets. *J. Am. Chem. Soc.* **1997**, *119* (45), 10903−10908.
(23) Pintado-Grima, C.; Bárcenas, O.; Bartolomé-Nafría, A.; Fornt-Suñé, M.; Iglesias, V.; Garcia-Pardo, J.; Ventura, S. A Review of Fifteen Years Developing Computational Tools to Study Protein Aggregation. *Biophysica* **2023**, *3* (1), 1.
(24) Iglesias, V.; Chilimoniuk, J.; Pintado-Grima, C.; Bárcenas, O.; Ventura, S.; Burdukiewicz, M. Aggregating Amyloid Resources: A Comprehensive Review of Databases on Amyloid-like Aggregation. *Comput. Struct. Biotechnol. J.* **2024**, *23*, 4011−4018.
(25) Taylor, A. I. P.; Staniforth, R. A. General Principles Underpinning Amyloid Structure. *Front. Neurosci.* **2022**, *16* (June), 878869.
(26) Cordeiro, Y.; Macedo, B.; Silva, J. L.; Gomes, M. P. B. Pathological Implications of Nucleic Acid Interactions with Proteins Associated with Neurodegenerative Diseases. *Biophys. Rev.* **2014**, *6* (1), 97−110.
(27) Chatani, E.; Yamamoto, N. Recent Progress on Understanding the Mechanisms of Amyloid Nucleation. *Biophys. Rev.* **2018**, *10* (2), 527−534.
(28) Fuertes, G.; Nevola, L.; Esteban-Martín, S. Perspectives on Drug Discovery Strategies Based on IDPs. In *Intrinsically Disordered Proteins*; Elsevier Inc., 2019; pp 275−327.
(29) Peña-Díaz, S.; Olsen, W. P.; Wang, H.; Otzen, D. E. Functional Amyloids: The Biomaterials of Tomorrow? *Adv. Mater.* **2024**, *36* (18), 2312823.
(30) Cui, H.; Yi, H.; Bao, H.; Tan, Y.; Tian, C.; Shi, X.; Gan, D.; Zhang, B.; Liang, W.; Chen, R.; Zhu, Q.; Fang, L.; Gao, X.; Huang, H.; Tian, R.; Sperling, S. R.; Hu, Y.; Chen, W. The SWI/SNF Chromatin Remodeling Factor DPF3 Regulates Metastasis of CcRCC by Modulating TGF-β Signaling. *Nat. Commun.* **2022**, *13* (1), 4680.
(31) Verrillo, G.; Obeid, A. M.; Genco, A.; Scrofani, J.; Orange, F.; Hanache, S.; Mignon, J.; Leyder, T.; Michaux, C.; Kempeneers, C.; Bricmont, N.; Herkenne, S.; Vernos, I.; Martin, M.; Mottet, D. Non-Canonical Role for the BAF Complex Subunit DPF3 in Mitosis and Ciliogenesis. *J. Cell Sci.* **2024**, *137* (9), jcs261744.
(32) Mignon, J.; Leyder, T.; Michaux, C. The Intrinsically Disordered DPF3 Zinc Finger Protein: A Promising New Target in Cancer Therapy. *J. Cancer Biol.* **2022**, *3* (2), 79−82.
(33) Mignon, J.; Mottet, D.; Leyder, T.; Uversky, V. N.; Perpète, E. A.; Michaux, C. Structural Characterisation of Amyloidogenic Intrinsically Disordered Zinc Finger Protein Isoforms DPF3b and DPF3a. *Int. J. Biol. Macromol.* **2022**, *218*, 57−71.
(34) Mignon, J.; Mottet, D.; Verrillo, G.; Matagne, A.; Perpète, E. A.; Michaux, C. Revealing Intrinsic Disorder and Aggregation Properties of the DPF3a Zinc Finger Protein. *ACS Omega* **2021**, *6* (29), 18793−18801.
(35) Leyder, T.; Mignon, J.; Mottet, D.; Michaux, C. Unveiling the Metal-Dependent Aggregation Properties of the C- Terminal Region of Amyloidogenic Intrinsically Disordered Protein Isoforms DPF3b and DPF3a C-Terminal Region of Amyloidogenic Intrinsically Disordered. *Int. J. Mol. Sci.* **2022**, *23* (23), 15291.
(36) Mignon, J.; Leyder, T.; Mottet, D.; Uversky, V. N.; Michaux, C. In-Depth Investigation of the Effect of PH on the Autofluorescence Properties of DPF3b and DPF3a Amyloid Fibrils. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2024**, *313*, 124156.
(37) Mignon, J.; Leyder, T.; Monari, A.; Mottet, D.; Michaux, C. Exploration of the Influence of Environmental Changes on the Conformational and Amyloidogenic Landscapes of the Zinc Finger Protein DPF3a by Combining Biophysical and Molecular Dynamics Approaches. *Int. J. Biol. Macromol.* **2025**, *310* (P1), 143234.
(38) Leyder, T.; Mignon, J.; Bongiovanni, E.; Machiels, Q.; Waeytens, J.; Raussens, V.; Monari, A.; Mottet, D.; Michaux, C. Unveiling the Effect of Phosphorylation and Phosphomimetics on the Structural and Aggregation Properties of the Amyloidogenic Intrinsically Disordered Protein DPF3a. *Int. J. Biol. Macromol.* **2026**, *335* (P2), 149393.
(39) Conchillo-Solé, O.; de Groot, N. S.; Avilés, F. X.; Vendrell, J.; Daura, X.; Ventura, S. AGGRESCAN: A Server for the Prediction and Evaluation of "Hot Spots" of Aggregation in Polypeptides. *BMC Bioinf.* **2007**, *8*, 65.
(40) Orlando, G.; Silva, A.; MacEdo-Ribeiro, S.; Raimondi, D.; Vranken, W. Accurate Prediction of Protein Beta-Aggregation with Generalized Statistical Potentials. *Bioinformatics* **2020**, *36* (7), 2076−2081.
(41) Burdukiewicz, M.; Sobczyk, P.; Rödiger, S.; Duda-Madej, A.; MacKiewicz, P.; Kotulska, M. Amyloidogenic Motifs Revealed by N-Gram Analysis. *Sci. Rep.* **2017**, *7* (1), 12961.
(42) Ahmed, A. B.; Znassi, N.; Château, M. T.; Kajava, A. V. A Structure-Based Approach to Predict Predisposition to Amyloidosis. *Alzheimer's Dementia* **2015**, *11* (6), 681−690.
(43) Sormanni, P.; Vendruscolo, M. Protein Solubility Predictions Using the Camsol Method in the Study of Protein Homeostasis. *Cold Spring Harb. Perspect. Biol.* **2019**, *11* (12), a033845.
(44) Gonay, V.; Dunne, M. P.; Caceres-Delpiano, J.; Kajava, A. V. Developing Machine-Learning-Based Amyloidogenicity Predictors with Cross-Beta DB. *Alzheimer's Dementia* **2025**, *21*, No. e14510.

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

(45) Tian, J.; Wu, N.; Guo, J.; Fan, Y. Prediction of Amyloid Fibril-Forming Segments Based on a Support Vector Machine. *BMC Bioinf.* **2009**, *10*, 8−12.
(46) Falgarone, T.; Villain, E.; Guettaf, A.; Leclercq, J.; Kajava, A. V. TAPASS: Tool for Annotation of Protein Amyloidogenicity in the Context of Other Structural States. *J. Struct. Biol.* **2022**, *214* (1), 107840.
(47) Maurer-Stroh, S.; Debulpaep, M.; Kuemmerer, N.; De La Paz, M. L.; Martins, I. C.; Reumers, J.; Morris, K. L.; Copland, A.; Serpell, L.; Serrano, L.; Schymkowitz, J. W. H.; Rousseau, F. Exploring the Sequence Determinants of Amyloid Structure Using Position-Specific Scoring Matrices. *Nat. Methods* **2010**, *7* (3), 237−242.
(48) Tsolis, A. C.; Papandreou, N. C.; Iconomidou, V. A.; Hamodrakas, S. J. A Consensus Method for the Prediction of "Aggregation-Prone" Peptides in Globular Proteins. *PLoS One* **2013**, *8* (1), No. e54175.
(49) Navarro, S.; Ventura, S. Computational Methods to Predict Protein Aggregation. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102343.
(50) Prabakaran, R.; Rawat, P.; Kumar, S.; Michael Gromiha, M.; AnuPP, M. A Versatile Tool to Predict Aggregation Nucleating Regions in Peptides and Proteins. *J. Mol. Biol.* **2021**, *433* (11), 166707.
(51) Keresztes, L.; Szögi, E.; Varga, B.; Farkas, V.; Perczel, A.; Grolmusz, V. The Budapest Amyloid Predictor and Its Applications. *Biomolecules* **2021**, *11* (4), 500.
(52) Gasior, P.; Kotulska, M. FISH Amyloid - a New Method for Finding Amyloidogenic Segments in Proteins Based on Site Specific Co-Occurrence of Aminoacids. *BMC Bioinf.* **2014**, *15* (1), 54.
(53) Garbuzynskiy, S. O.; Lobanov, M. Y.; Galzitskaya, O. V. FoldAmyloid: A Method of Prediction of Amyloidogenic Regions from Protein Sequence. *Bioinformatics* **2010**, *26* (3), 326−332.
(54) Thangakani, A. M.; Kumar, S.; Nagarajan, R.; Velmurugan, D.; Gromiha, M. M. GAP: Towards Almost 100% Prediction for β-Strand-Mediated Aggregating Peptides with Distinct Morphologies. *Bioinformatics* **2014**, *30*, 1983−1990.
(55) Emily, M.; Talvas, A.; Delamarche, C. MetAmyl: A METa-Predictor for AMYLoid Proteins. *PLoS One* **2013**, *8* (11), No. e79722.
(56) Thompson, M. J.; Sievers, S. A.; Karanicolas, J.; Ivanova, M. I.; Baker, D.; Eisenberg, D. The 3D Profile Method for Identifying Fibril-Forming Segments of Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (11), 4074−4078.
(57) Delgado, J.; Radusky, L. G.; Cianferoni, D.; Serrano, L. FoldX 5.0: Working with RNA, Small Molecules and a New Graphical Interface. *Bioinformatics* **2019**, *35* (20), 4168−4169.
(58) Louros, N.; Orlando, G.; De Vleeschouwer, M.; Rousseau, F.; Schymkowitz, J. Structure-Based Machine-Guided Mapping of Amyloid Sequence Space Reveals Uncharted Sequence Clusters with Higher Solubilities. *Nat. Commun.* **2020**, *11* (1), 3314.
(59) Ghomi, H. T.; Thompson, J. J.; Lill, M. A. Are Distance-Dependent Statistical Potentials Considering Three Interacting Bodies Superior to Two-Body Statistical Potentials for Protein Structure Prediction? *J. Bioinform. Comput. Biol.* **2014**, *12* (05), 1450022.
(60) Tabatabaei Ghomi, H.; Topp, E. M.; Lill, M. A. Fibpredictor: A Computational Method for Rapid Prediction of Amyloid Fibril Structures. *J. Mol. Model.* **2016**, *22* (9), 206.
(61) Shen, M.; Sali, A. Statistical Potential for Assessment and Prediction of Protein Structures. *Protein Sci.* **2006**, *15* (11), 2507−2524.
(62) Wojciechowski, J. W.; Kotulska, M. PATH - Prediction of Amyloidogenicity by Threading and Machine Learning. *Sci. Rep.* **2020**, *10* (1), 7721.
(63) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859−1865.
(64) Park, S.-J.; Kern, N.; Brown, T.; Lee, J.; Im, W. C. H. A. R. M. M.-G. U. I. PDB Manipulator: Various PDB Structural Modifications for Biomolecular Modeling and Simulation. *J. Mol. Biol.* **2023**, *435* (14), 167995.
(65) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12* (1), 405−413.
(66) Yuan, S.; Chan, H. C. S.; Hu, Z. Using PyMOL as a Platform for Computational Drug Design. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7* (2), No. e1298.
(67) Kern, N. R.; Lee, J.; Choi, Y. K.; Im, W. CHARMM-GUI Multicomponent Assembler for Modeling and Simulation of Complex Multicomponent Systems. *Nat. Commun.* **2024**, *15* (1), 5459.
(68) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701−1718.
(69) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1-2*, 19−25.
(70) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71−73.
(71) Samantray, S.; Yin, F.; Kav, B.; Strodel, B. Different Force Fields Give Rise to Different Amyloid Aggregation Pathways in Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2020**, *60* (12), 6462−6475.
(72) Man, V. H.; He, X.; Gao, J.; Wang, J. Effects of All-Atom Molecular Mechanics Force Fields on Amyloid Peptide Assembly: The Case of PHF6 Peptide of Tau Protein. *J. Chem. Theory Comput.* **2021**, *17* (10), 6458−6471.
(73) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864−1874.
(74) Gao, Y.; Lee, J.; Smith, I. P. S.; Lee, H.; Kim, S.; Qi, Y.; Klauda, J. B.; Widmalm, G.; Khalid, S.; Im, W. C. H. A. R. M. M.-G. U. I. Supports Hydrogen Mass Repartitioning and Different Protonation States of Phosphates in Lipopolysaccharides. *J. Chem. Inf. Model.* **2021**, *61* (2), 831−839.
(75) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463−1472.
(76) Müller, K.; Brown, L. D. Location of Saddle Points and Minimum Energy Paths by a Constrained Simplex Optimization Procedure. *Theor. Chim. Acta* **1979**, *53* (1), 75−93.
(77) Van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-Frog Algorithm for Stochastic Dynamics. *Mol. Simul.* **1988**, *1* (3), 173−185.
(78) Bernetti, M.; Bussi, G. Pressure Control Using Stochastic Cell Rescaling. *J. Chem. Phys.* **2020**, *153* (11), 114107.
(79) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.
(80) Petersen, H. G. Accuracy and Efficiency of the Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (9), 3668−3679.
(81) Andersen, C. A. F.; Palmer, A. G.; Brunak, S.; Rost, B. Continuum Secondary Structure Captures Protein Flexibility. *Structure* **2002**, *10* (2), 175−184.
(82) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.
(83) Wiegand, T.; Malär, A. A.; Cadalbert, R.; Ernst, M.; Böckmann, A.; Meier, B. H. Asparagine and Glutamine Side-Chains and Ladders in HET-s(218−289) Amyloid Fibrils Studied by Fast Magic-Angle Spinning NMR. *Front. Mol. Biosci.* **2020**, *7*, 582033.
(84) Gazit, E. A Possible Role for Π-stacking in the Self-assembly of Amyloid Fibrils. *FASEB J.* **2002**, *16* (1), 77−83.
(85) Makin, O. S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L. C. Molecular Basis for Amyloid Fibril Formation and Stability. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (2), 315−320.

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494

(86) Richaud, A. D.; Mandal, S.; Das, A.; Roche, S. P. Tunable CH/π Interactions within a Tryptophan Zipper Motif to Stabilize the Fold of Long β-Hairpin Peptides. *ACS Chem. Biol.* **2023**, *18*, 2555−2563.
(87) Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. C-H···π-Interactions in Proteins. *J. Mol. Biol.* **2001**, *307*, 357−377.
(88) Chelli, R.; Gervasio, F. L.; Procacci, P.; Schettino, V. Stacking and T-Shape Competition in Aromatic-Aromatic Amino Acid Interactions. *J. Am. Chem. Soc.* **2002**, *124*, 6133−6143.
(89) Kurt, T. D.; Aguilar-Calvo, P.; Jiang, L.; Rodriguez, J. A.; Alderson, N.; Eisenberg, D. S.; Sigurdson, C. J. Asparagine and Glutamine Ladders Promote Cross-Species Prion Conversion. *J. Biol. Chem.* **2017**, *292* (46), 19076−19086.
(90) Yamazaki, M.; Ikeda, K.; Kameda, T.; Nakao, H.; Nakano, M. Kinetic Mechanism of Amyloid-β-(16−22) Peptide Fibrillation. *J. Phys. Chem. Lett.* **2022**, *13* (26), 6031−6036.
(91) Mayes, M. L.; Perreault, L. Probing the Nature of Noncovalent Interactions in Dimers of Linear Tyrosine-Based Dipeptides. *ACS Omega* **2019**, *4* (1), 911−919.
(92) Wilkinson, M.; Xu, Y.; Thacker, D.; Taylor, A. I. P.; Fisher, D. G.; Gallardo, R. U.; Radford, S. E.; Ranson, N. A. Structural Evolution of Fibril Polymorphs during Amyloid Assembly. *Cell* **2023**, *186* (26), 5798−5811.
(93) Lutter, L.; Aubrey, L. D.; Xue, W. F. On the Structural Diversity and Individuality of Polymorphic Amyloid Protein Assemblies. *J. Mol. Biol.* **2021**, *433*, 167124.
(94) Pretti, E.; Shell, M. S. Mapping the Configurational Landscape and Aggregation Phase Behavior of the Tau Protein Fragment PHF6. *Proc. Natl. Acad. Sci. U. S. A.* **2023**, *120* (48), No. e2309995120.
(95) Deike, S.; Rothemund, S.; Voigt, B.; Samantray, S.; Strodel, B.; Binder, W. H. β-Turn Mimetic Synthetic Peptides as Amyloid-β Aggregation Inhibitors. *Bioorg. Chem.* **2020**, *101*, 104012.
(96) Cerf, E.; Sarroukh, R.; Tamamizu-Kato, S.; Breydo, L.; Derclaye, S.; Dufrêne, Y. F.; Narayanaswami, V.; Goormaghtigh, E.; Ruysschaert, J.-M.; Raussens, V. Antiparallel β-Sheet: A Signature Structure of the Oligomeric Amyloid β-Peptide. *Biochem. J.* **2009**, *421* (3), 415−423.
(97) Zanjani, A. A. H.; Reynolds, N. P.; Zhang, A.; Schilling, T.; Mezzenga, R.; Berryman, J. T. Amyloid Evolution: Antiparallel Replaced by Parallel. *Biophys. J.* **2020**, *118* (10), 2526−2536.
(98) Zhaliazka, K.; Kurouski, D. Nanoscale Characterization of Parallel and Antiparallel β-Sheet Amyloid Beta 1−42 Aggregates. *ACS Chem. Neurosci.* **2022**, *13* (19), 2813−2820.
(99) Cehlar, O.; Njemoga, S.; Horvath, M.; Cizmazia, E.; Bednarikova, Z.; Barrera, E. E. Structures of Oligomeric States of Tau Protein, Amyloid-β, α-Synuclein and Prion Protein Implicated in Alzheimer's Disease, Parkinson's Disease and Prionopathies. *Int. J. Mol. Sci.* **2024**, *25* (23), 13049.
(100) Liu, H.; Zhong, H.; Liu, X.; Zhou, S.; Tan, S.; Liu, H.; Yao, X. Disclosing the Mechanism of Spontaneous Aggregation and Template-Induced Misfolding of the Key Hexapeptide (PHF6) of Tau Protein Based on Molecular Dynamics Simulation. *ACS Chem. Neurosci.* **2019**, *10* (12), 4810−4823.
(101) Ganguly, P.; Do, T. D.; Larini, L.; Lapointe, N. E.; Sercel, A. J.; Shade, M. F.; Feinstein, S. C.; Bowers, M. T.; Shea, J. E. Tau Assembly: The Dominant Role of PHF6 (VQIVYK) in Microtubule Binding Region Repeat R3. *J. Phys. Chem. B* **2015**, *119* (13), 4582−4593.
(102) Wang, J.; Liu, Z.; Zhao, S.; Zhang, Y.; Xu, T.; Li, S. Z.; Li, W. Aggregation Rules of Short Peptides. *JACS Au* **2024**, *4* (9), 3567−3580.
(103) Tsai, H. H.; Reches, M.; Tsai, C. J.; Gunasekaran, K.; Gazit, E.; Nussinov, R. Energy Landscape of Amyloidogenic Peptide Oligomerization by Parallel-Tempering Molecular Dynamics Simulation: Significant Role of Asn Ladder. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (23), 8174−8179.
(104) Wojciechowski, J. W.; Szczurek, W.; Szulc, N.; Szefczyk, M.; Kotulska, M. PACT - Prediction of Amyloid Cross-Interaction by Threading. *Sci. Rep.* **2023**, *13* (1), 22268.
(105) Koliński, M.; Dec, R.; Dzwolak, W. Multiscale Modeling of Amyloid Fibrils Formed by Aggregating Peptides Derived from the Amyloidogenic Fragment of the A-Chain of Insulin. *Int. J. Mol. Sci.* **2021**, *22* (22), 12325.
(106) Puławski, W.; Koliński, A.; Koliński, M. Multiscale Modeling of Protofilament Structures: A Case Study on Insulin Amyloid Aggregates. *Int. J. Biol. Macromol.* **2025**, *285*, 138382.
(107) Samantray, S.; Schumann, W.; Illig, A.-M.; Carballo-Pacheco, M.; Paul, A.; Barz, B.; Strodel, B. Molecular Dynamics Simulations of Protein Aggregation: Protocols for Simulation Setup and Analysis with Markov State Models and Transition Networks. In *Computer Simulations of Aggregation of Proteins and Peptides*; Li, M. S., Kloczkowski, A., Cieplak, M., Kouza, M., Eds.; Springer US: New York, NY, 2022; pp 235−279.
(108) Das, M.; Bhargava, B. L. Exploring the Candidates for a New Protein Folding-Cross-α Amyloid-in Available Protein Databases. *Phys. Chem. Chem. Phys.* **2020**, *22*, 23725−23734.
(109) Ragonis-Bachar, P.; Axel, G.; Blau, S.; Ben-Tal, N.; Kolodny, R.; Landau, M. What Can AlphaFold Do for Antimicrobial Amyloids? *Proteins: Struct., Funct., Bioinf.* **2024**, *92* (2), 265−281.

https://doi.org/10.1021/acs.jcim.6c00951
J. Chem. Inf. Model. 2026, 66, 5471−5494