PNAS nexus

PNAS Nexus, 2022, **1**, 1–12
https://doi.org/10.1093/pnasnexus/pgac263
Advance access publication date: 25 November 2022
Research Report

Sequence patterns and signatures: Computational and experimental discovery of amyloid-forming peptides

Xingqing Xiao [a,1], Alicia S. Robang[b,1], Sudeep Sarma [a,1], Justin V. Le[b], Michael E. Helmicki[b], Matthew J. Lambert[b], Ricardo Guerrero-Ferreira[c], Johana Arboleda-Echavarria[c], Anant K. Paravastu [b,*] and Carol K. Hall [a,*]

[a]Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC 27695, USA
[b]Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[c]Robert P. Apkarian Integrated Electron Microscopy Core, Emory University School of Medicine, Atlanta, GA 30322, USA
*To whom correspondence should be addressed. Email: hall@ncsu.edu, anant.paravastu@chbe.gatech.edu
Edited By: Shibu Yooseph

Abstract

Screening amino acid sequence space via experiments to discover peptides that self-assemble into amyloid fibrils is challenging. We have developed a computational *pep*tide *a*ssembly *d*esign (PepAD) algorithm that enables the discovery of amyloid-forming peptides. Discontinuous molecular dynamics (DMD) simulation with the PRIME20 force field combined with the FoldAmyloid tool is used to examine the fibrilization kinetics of PepAD-generated peptides. PepAD screening of ∼10,000 7-mer peptides resulted in twelve top-scoring peptides with two distinct hydration properties. Our studies revealed that eight of the twelve *in silico* discovered peptides spontaneously form amyloid fibrils in the DMD simulations and that all eight have at least five residues that the FoldAmyloid tool classifies as being aggregation-prone. Based on these observations, we re-examined the PepAD-generated peptides in the sequence pool returned by PepAD and extracted five sequence patterns as well as associated sequence signatures for the 7-mer amyloid-forming peptides. Experimental results from Fourier transform infrared spectroscopy (FTIR), thioflavin T (ThT) fluorescence, circular dichroism (CD), and transmission electron microscopy (TEM) indicate that all the peptides predicted to assemble *in silico* assemble into antiparallel β-sheet nanofibers in a concentration-dependent manner. This is the first attempt to use a computational approach to search for amyloid-forming peptides based on customized settings. Our efforts facilitate the identification of β-sheet-based self-assembling peptides, and contribute insights towards answering a fundamental scientific question: "*What does it take, sequence-wise, for a peptide to self-assemble?*"

Keywords: peptide assembly design, discontinuous molecular dynamics, amyloid-forming peptides, fourier-transform infrared spectroscopy, transmission electron microscopy

Significance Statement:

Self-assembling peptides that can form different supramolecular architectures can be useful in various drug delivery and biomedical engineering applications. The discovery of peptides that can self-assemble to form amyloid-like structures lacks a systematic approach and there has been limited work by researchers to search for peptide sequences that are predicted to form cross-β spines. In this paper, we describe a novel, state-of-the-art computational approach that combines a *Pep*tide Assembly Design (PepAD) algorithm and discontinuous molecular dynamics simulations of amyloid formation to predict peptide sequences and patterns that can self-assemble to form cross-β spines. All eight of the discovered peptide sequences assemble to form antiparallel β-sheets in experimental measurements.

Introduction

Peptide self-assembly is a process in which peptides spontaneously form ordered aggregates. Peptide self-assembly into nanoscale architectures (1–3) provides numerous advantages for applications, including drug release (4, 5), protein scaffolds (6, 7), tissue cell culture (8), and biomimetic 3D printing (9). Rational design of peptides from an alphabet of 20 natural amino acids makes for a huge number of possible sequences, yet only a handful of peptides have been discovered that self-assemble to form particular tertiary structures with unique bio-functions (10, 11). Great structural variety can be achieved, in principle, via a "bottom-up" strategy in which many factors, such as amino acid composition, sequence length, and pattern, are tailored to form desired functional architectures (12, 13). Short peptides (≤ 10 residues) are of interest in bioengineering because they can be efficiently synthesized, greatly reduce the level of complexity in sequence design, and make it easier to establish connections between peptide sequence and aggregate structure. For instance, Frederix et al. (14) conducted a comprehensive simulation study to examine the self-assembly behaviors of all possible tripeptides ($20^3 = 8{,}000$). Only

[1]**Equal Contribution:** X.X, A.S.R. and S.S. contributed equally to the manuscript.
[2]**Competing interest:** The authors declare no conflict of interest.
Received: April 19, 2022. **Accepted:** November 22, 2022

thirteen of these tripeptides were found to form a well-ordered hydrogel network, supporting conventional wisdom that finding new sequences that self-assemble is challenging but not impossible.

The abnormal misfolding of peptide fragments and subsequent self-assembly into amyloid fibrils is correlated with many neurodegenerative diseases (15, 16), including Alzheimer's, Parkinson's, and Huntington's diseases. Amyloid fibrils are 1D β-sheet-rich peptide aggregates. Thus far, most of the known β-sheet-forming peptides are derived from natural proteins (17–19). For example, the 7-mer peptide fragment Aβ (16–22) (sequence: *KLVFFAE*), which is associated with Alzheimer's disease, can assemble into β-sheet fibrils. A well-established motif for human-designed β-sheet-forming peptides is a pattern of alternating hydrophilic and polar residues along the peptide chain; these have been found experimentally to form multilayer β-sheet structures with hydrophobic interfaces between the layers (20, 21). Sawaya and Eisenberg proposed eight possible classes of cross-β spines consisting of a 2-layer β-sheet and based on their work researchers have been trying to develop methods to search for amyloid-forming peptides in nature that self-assemble into these cross-β spines (22). To date, some promising bioinformatic methods have been proposed to predict the amyloidogenic region of a given peptide sequence (23–25). For example, TANGO is a statistical mechanics algorithm that predicts protein aggregation (26, 27). It is based on the assumption that in -aggregates, the "aggregation-prone" residues within a protein are fully buried and tend to have their hydrogen-bonding potential satisfied. Fernandez-Escamilla et al. used TANGO to examine the aggregation of a set of 179 literature-reported peptides and 71 disease-related peptides. They found that while TANGO can accurately predict the self-aggregation propensity of short peptides (less than 20 residues), it has difficulty predicting that of long peptides (more than 30 residues) (26). Although such bioinformatic methods are powerful amyloidogenicity predicting tools, they cannot be applied to discover, which peptide sequences can form amyloid-like structures in a fast and automatic way. To address this issue, we are developing a computational toolkit that enables custom presettings, such as peptide lengths, arrangements, assembly scaffolds, and hydration properties, to identify fibril-forming peptides.

In previous work, we developed a computational *pep*tide *b*inding *d*esign (PepBD) algorithm to discover peptide-based binders for biomolecular targets (28–32), e.g., proteins and RNA. The PepBD algorithm can be used to design peptides with distinct hydration properties tailored to an experimental researcher's needs. The algorithm uses atomistic force fields rather than knowledge-based information for the peptide designs, so it enables the discovery of high-affinity binding peptides to targets that have no known binders available in the protein data bank. The PepBD algorithm has been used successfully to design 15-mer transfer RNA^{Lys3}-binding peptides for the inhibition of HIV reverse transcriptase (33), 12-mer peptide-based biological recognition elements for cardiac troponin I (34), and for neuropeptide Y (35) to detect human performance indicators, peptide inhibitors targeting *Clostridioides difficile* toxins to neutralize the cytopathic effects of the toxins (36), Protein A/L mimetic peptide ligands that bind to immunoglobulin G for monoclonal antibody purification (37, 38) peptides that bind to the Receptor Binding Domain of the SARS-CoV-2 spike-protein (39).

In this work, a *pep*tide *a*ssembly *d*esign (PepAD) algorithm is described to discover self-assembling peptides for the construction of particular supramolecular structures, as opposed to discovering binding peptides for the recognition of specific biomolecular targets as is done in PepBD. The goal of this project is to discover peptides (not known in nature) that self-assemble to form a particular type of amyloid-like structure, specifically the Class-8 cross-β spine described by Sawaya et al. (22) This structure contains a 2-layer β-sheet structure, with antiparallel-oriented β-strands in each layer and parallel-oriented β-strands between two layers. The PepAD algorithm is employed to search for 7-mer peptides that self-assemble spontaneously to form the cross-β spine of the 8th class, thereby leading to a pool of candidate peptide sequences. PepAD screening provided approximately 10,000 distinct peptide sequences that could self-assemble to the desired structure. Out of these ~10,000 peptides, the top twelve of the PepAD-generated *in silico* peptides, viz. P1-P12, were evaluated using the FoldAmyloid tool. Discontinuous molecular dynamics (DMD) simulations with the PRIME20 force field were carried out to examine the fibrillization kinetics of the twelve *in silico* peptides at a temperature of 296.1 K. Our computational studies revealed that eight of the 12 *in silico* discovered peptides, viz. P1, P2, P5, P7, P9, P10, P11, and P12, spontaneously form amyloid fibrils in the DMD/PRIME20 simulations, and that all of these peptides have at least five residues along the chain with self-aggregation scales higher than a threshold value as classified by the FoldAmyloid tool. Based on this bioinformatic information, we examine 300 candidate peptide sequences screened by PepAD and identify approximately 240 peptide sequences that have at least five residues along the chain with self-aggregation scales higher than a threshold value using FoldAmyloid. The characterization of the 240 promising *in silico* peptides leads to five sequence patterns as well as to their associated sequence signatures for 7-mer amyloid-forming peptides. A sequence pattern is an arrangement of *h*ydrophobic (H), *c*harged (C), and *p*olar (P) residues along the chain. The sequence signature is a subset of this pattern, a family of sequences with specific amino acids at certain key sites and a broader choice of amino acids at other sites along the chain. Finally, we experimentally test the eight peptides predicted by DMD/PRIME20 to form amyloid-like structures and discovered that they assemble into antiparallel β-sheets at concentrations between 0.2 and 10 mM at room temperature in water. The entire strategy of discovering peptides that self-assemble to form amyloid-like structures integrates computational design (PepAD), simulation-based screening (DMD) and experimental testing (see Fig. 1). Experimental tests for fibril formation includes Thioflavin T (ThT) fluorescence, circular dichroism (CD), Fourier transform infrared spectroscopy (FTIR), and transmission electron microscopy (TEM).

Results

Computational discovery of amyloid-forming peptides

A PepAD algorithm based on the Monte Carlo (MC) algorithm has been developed in our lab to design peptide sequences that can self-aggregate to form supramolecular architectures. The designed sequences are draped upon a fixed peptide backbone scaffold within a two-layer fibril structure; this structure is fixed throughout the design process. Two kinds of sequence change moves are applied: residue mutation and residue exchange, to search for new self-assembling peptide candidates. A score function, Γ_{score}, which considers (i) the binding free energy, $\Delta G_{binding}$, of the peptide chain with its neighboring peptides (40, 41) and (ii) the intrinsic self-aggregation propensity, $P_{aggregation}$, of the individual peptides (42–44), is used to evaluate the acceptance of new peptide candidates. $P_{aggregation}$ evaluates the aggregation propen-

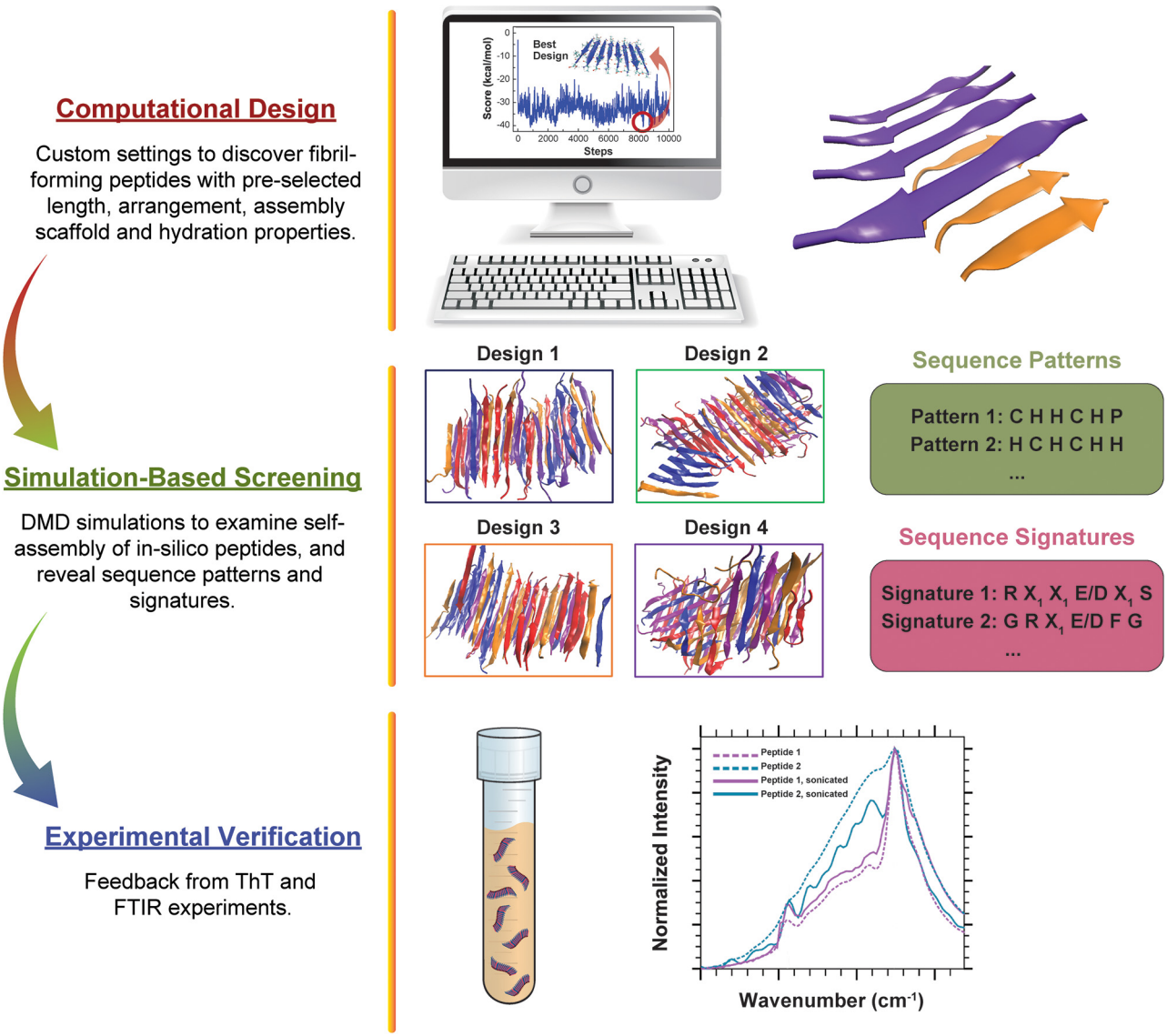

Fig. 1. An integrated strategy containing customized computational design, simulation-based screening, and experimental verification is described to discover amyloid-forming peptides.

sity of the fibril-like supramolecular structure. Hence, the score function Γ_{score} is defined to be

$$\Gamma_{score} = \Delta G_{binding} - \lambda \times P_{aggregation}, \quad (1)$$

where λ is a weighting factor that adjusts the relative importance of the intrinsic aggregation propensity of the peptides during the sequence evolution. The $\Delta G_{binding}$ is defined as

$$\Delta G_{binding} = \sum \left(G^{ij} - G^{i} - G^{j} \right) . \quad (2)$$

The notation i and j is used to represent two neighboring peptides, respectively. The G^{ij} indicates the free energy of the pair of peptides i and j in solution, and the G^{i} and G^{j} indicate the free energies of each individual peptide i and j, respectively. An all-atom amino acid model is used, and the force field parameters are taken from the Amber 14SB force field. The $P_{aggregation}$ is defined as

$$P_{aggregation} = \delta_{hydr} I_{hydr} + \delta_{\alpha} I_{\alpha} + \delta_{\beta} I_{\beta} + \delta_{ch} I_{ch} + \delta_{pat} I_{pat}, \quad (3)$$

where I_{hydr} indicates the hydrophobicity of the sequence, I_{α} is the α-helical propensity, I_{β} is the -sheet propensity, I_{ch} is the absolute value of the net charge of the sequence, and I_{pat} is the sequence pattern that accounts for the sequence-dependent nature of peptide secondary structure. The coefficients δ_i are used to weigh each of the individual factors. The algorithm allows for the design of peptides at distinct pH environments by considering the protonation and deprotonation state of the natural amino acids at the system pH. A description of the computational design procedure and Γ_{score} is provided in "PepAD algorithm" and the Supplementary Material.

As explained above, an initial peptide backbone scaffold is required in the PepAD algorithm to discover potential sequences that are capable of assembling into the Class-8 cross-β spine (22). The Aβ (16–22) (sequence: KLVFFAE) was used because it is known to form the characteristic Class-8 structure, a 2-layer amyloid fibril that consists of antiparallel-oriented β-strands in each layer, and parallel-oriented β-strands between two layers (45). The structure of Aβ (16–22) is examined in DMD simulation, which predicts that the sequence forms amyloid fibrils (Fig. 2A), consistent

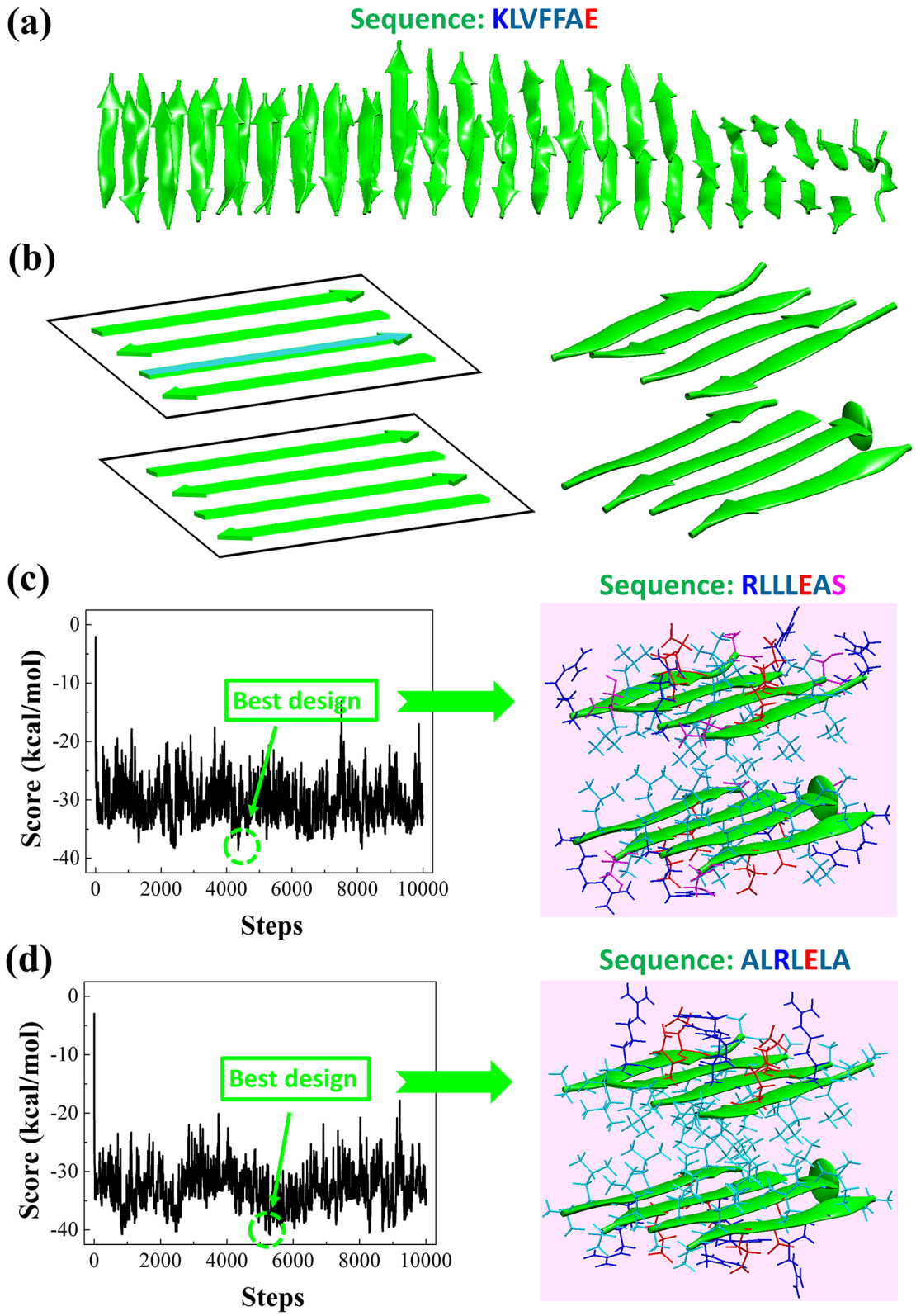

Fig. 2. (a) The fibril formed by the Aβ (16–22) peptide *KLVFFAE* in the DMD simulations. (b) Input structure for PepAD: A 2-layer β-sheet structure is constructed using the Discovery Studio 3.5 package and optimized by atomistic MD simulation in the AMBER14 package; the 8 Aβ (16–22) peptides are aligned in antiparallel arrangements within each β-sheet layer and in parallel arrangements between the β-sheet layers. Sequence evolution on the 2-layer β-sheet structure leads to the fluctuation of scores in the PepAD algorithm, resulting in the two top-scoring 7-mer peptides: (c) sequence RLLLEAS in case 1 and (d) sequence ALRLELA in case 2.

Table 1. Sequences of twelve top-scoring *in silico* peptides, and their associated Γ_{score}, $\Delta G_{binding}$, and $\lambda \times P_{aggregation}$ values. (unit: kcal/mol.)

Case 1										
Peptides	**Sequences**							Γ_{score}	$\Delta G_{binding}$	$\lambda \times P_{aggregation}$
P1[(a)]	R	L	L	L	E	A	S	−38.59	−28.08	10.51
P2	R	A	L	L	S	V	E	−36.74	−26.64	10.09
P3	I	G	R	A	E	W	S	−37.52	−29.50	8.02
P4[(b)]	E	W	L	A	Q	A	R	−41.76	−25.88	15.88
P5	A	L	R	L	S	L	E	−42.78	−25.84	16.93
P6	Q	A	R	A	E	W	L	−43.54	−26.50	17.05
Case 2										
P7[(c)]	R	L	L	A	A	L	E	−36.79	−27.67	9.12
P8	L	W	R	L	E	G	G	−35.60	−28.17	7.43
P9	R	W	L	A	V	A	E	−35.82	−29.64	6.18
P10[(d)]	I	A	R	L	E	L	A	−40.19	−27.43	12.76
P11	R	A	L	L	A	L	E	−40.89	−27.22	13.67
P12	A	L	R	L	E	L	A	−41.81	−26.97	14.84

Peptides (a) P1—P3 and (b) P4—P6 were evolved in case 1 at $\lambda = 2.0$ and at $\lambda = 3.0$, respectively, and peptides (c) P7—P9 and (d) P10—P12 were evolved in case 2 at $\lambda = 2.0$ and at $\lambda = 3.0$, respectively.

with experimental findings. Based on the DMD prediction of Aβ (16–22)'s fibril formation, we construct a 2-layer β-sheet structure in the Discovery Studio 3.5 package. Eight 7-mer Aβ (16–22) peptides are aligned in an antiparallel arrangement within each β-sheet layer (4 peptides per layer) and in a parallel arrangement between the β-sheet layers. Later, a 5-ns explicit-solvent atomistic molecular dynamics simulation is conducted using the AMBER14 package to relax the 2-layer β-sheet structure to eliminate unfavorable atomic overlaps. This optimized 2-layer β-sheet structure is used as the initial backbone scaffold in the PepAD algorithm (Fig. 2b).

Through presetting two sets of hydration properties: case 1 with ($N_{hydrophobic} = 4$, $N_{polar} = 1$, $N_{charge} = 2$, and $N_{other} = 0$), and case 2 with ($N_{hydrophobic} = 5$, $N_{polar} = 0$, $N_{charge} = 2$, and $N_{other} = 0$), we search for amyloid-forming peptides via the PepAD algorithm. It is of interest to determine which hydration properties favor the amyloid-β type fibrilization of the peptides. If too many hydrophobic residues appear on the chain, the peptide is likely to be insoluble in solution, but if too many polar residues appear on the chain, the peptide is not likely to aggregate in solution. Two values are set for the weighting factor λ, viz. $\lambda = 2.0$ and $\lambda = 3.0$ in Equation (1) to avoid local searches during the evolution process. The pH value was set to be between 6.0 and 8.3. The weighting factors $\lambda = 2.0$ and $\lambda = 3.0$ were chosen to provide a good balance between optimizing the binding free energy and the aggregation propensity in the Γ_{score} of the amyloid-forming structure. All the searches start with random peptide sequences draped on the fixed backbone scaffold. Fig. 2c and d show two examples of Γ_{score} vs. the number of steps for the two cases. A moderate fluctuation in the score profiles is observed consistently in the entire design process as the residue(s) are mutated and exchanged to search for the best peptide candidates in a broad sequence space. The lowest scores in the profile indicate good peptide designs: the best sequence, *RLLLEAS*, in case 1 has the lowest $\Gamma_{score} = -38.59$ kcal/mol at $\lambda = 2.0$ (Fig. 2c), and the best sequence, *ALRLELA*, in case 2, has the lowest $\Gamma_{score} = -41.81$ kcal/mol at $\lambda = 3.0$ (Fig. 2d).

By ranking the scores of the evolved peptides, we identify a total of 300 potential amyloid-forming peptides: 150 for case 1 and 150 for case 2. Table 1 lists the twelve top-scoring *in silico* peptide sequences, as well as their associated scores (Γ_{score}), binding free energies ($\Delta G_{binding}$), and self-aggregation propensities ($\lambda \times P_{aggregation}$) calculated using Equations (1–3). For convenience, we call the twelve top-scoring peptides "P1" to "P12." peptides (a) P1—P3 and (b) P4—P6 are evolved in case 1 at $\lambda = 2.0$ and $\lambda = 3.0$, respectively, and peptides (c) P7—P9 and (d) P10—P12 are evolved in case 2 at $\lambda = 2.0$ and $\lambda = 3.0$, respectively.

Computational analysis of the fibrillization behaviors of the *in silico* peptides

A bioinformatics method, FoldAmyloid, is utilized to characterize peptide designs that result from the PepAD algorithm. In this method, peptides are predicted to be amyloidogenic if they contain at least seven consecutive residues that have average self-aggregation scales that are higher than an empirical threshold value of 21.4; otherwise, they are predicted to be non-amyloidogenic (46). A detailed discussion of the FoldAmyloid server is provided in the Supplementary Materials. The intrinsic amyloid propensities of the 12 PepAD-generated *in silico* peptides are evaluated in the FoldAmyloid web server. This method predicts the self-aggregation scales for the segments on the *in silico* peptides P1—P12, giving insight to which residues are aggregation-prone. These data are shown in Figs. 3 and S2. Eight peptides: P1, P2, P5, P7, P9, P10, P11, and P12 are found to have at least five residues along the chains with self-aggregation scales above 21.4. For instance, the residues (*R*, *L*, *L*, *L*, *E*) at sites 1–5 on peptide P1 are aggregation prone as their self-aggregation scales are higher than 21.4 (Fig. 3a). In contrast, four of the peptides: P3, P4, P6, and P8 are found to have less than five residues with self-aggregation scales above 21.4. For example, only one residue, *R*, at site three on P3 has a high self-aggregation scale value of 21.6 (Fig. 3b).

Next, we investigate the fibrillization kinetics of the 12 *in silico* peptides at 296.1 K using DMD/PRIME20 simulations. DMD/PRIME20 is a fast alternative to traditional molecular dynamics simulations that uses discontinuous potentials to model proteins. The force field PRIME20, developed by the Hall group in 2010, is arguably the most detailed and realistic of the coarse-grained models used to study protein aggregation. Each amino acid is represented by a 3-sphere backbone comprised of the united atoms NH, $C_\alpha H$, and CO and a single-sphere side chain, R. Water is modeled implicitly (47–52). Details are provided in "PepAD algorithm". For each *in silico* peptide system, we placed 48 identical peptide chains in a simulation box (200.0, 200.0, and 200.0 Å) to achieve a concentration of ~10 mM. The systems were simulated for 5 μs to examine their self-aggregation kinetics. It is worth not-

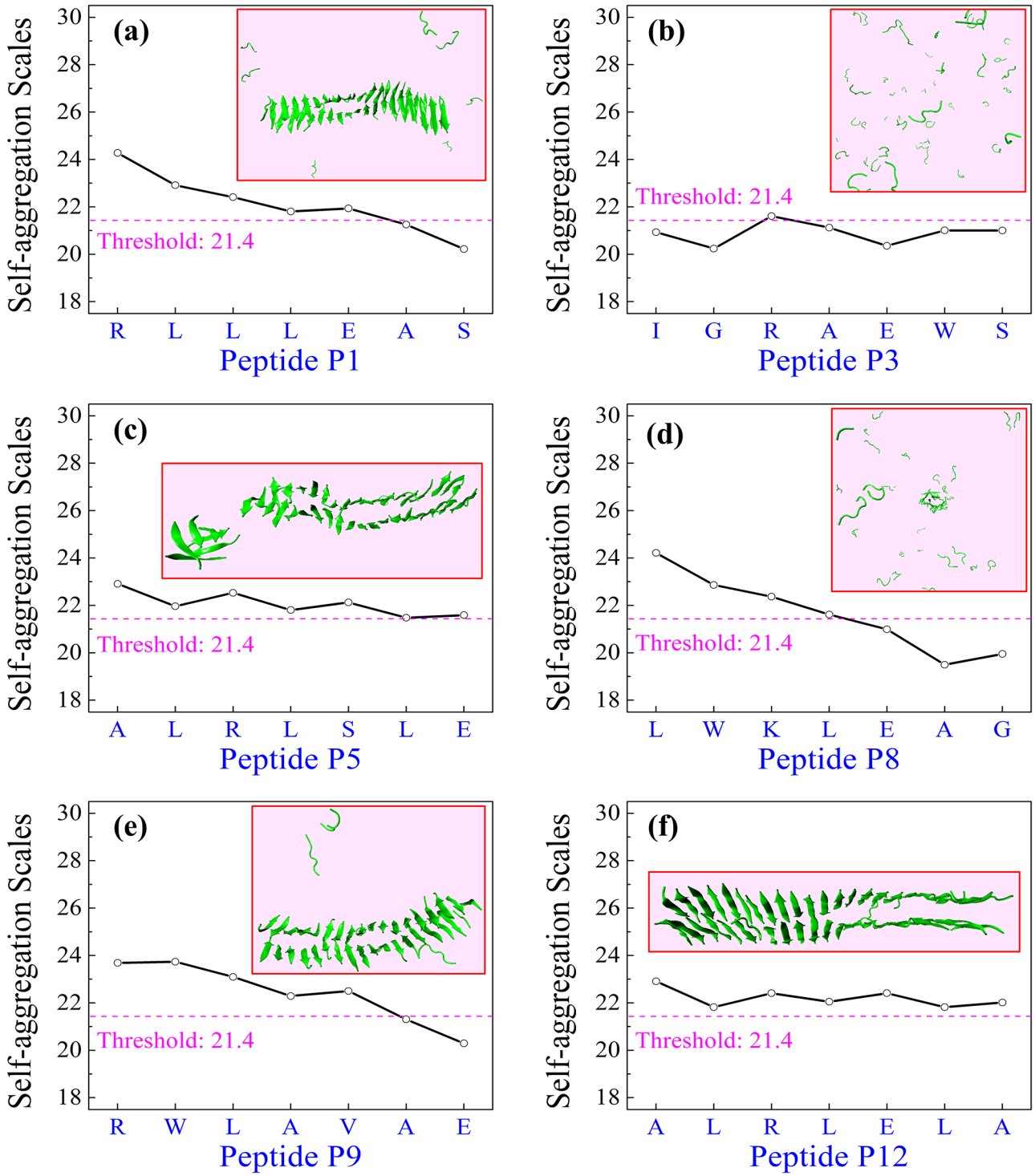

Fig. 3. The self-aggregation propensities of the PepAD-generated peptides (a) P1, (b) P3, (c) P5, (d) P8, (e) P9, and (f) P12 are evaluated using the FoldAmyloid tool (45). The final structures of the six *in silico* peptides in the DMD/PRIME20 simulations are inserted as insets to illustrate their fibrillization behaviors. Each simulation contains 48 peptides at a concentration of ~10 mM and a temperature of 296.1 K.

ing that all of the 48 peptide chains in the DMD simulations are in completely random configurations at the beginning. The simulated amyloid-like structures of the peptides P1—P12 are shown as insets in Figs. 3 and S2. Our simulation predictions agree with the results predicted by the FoldAmyloid server. All of the eight peptides that have at least five residues along the chains with self-aggregation scales above 21.4: P1 (Fig. 3a), P2 (Fig. S2a), P5 (Fig. 3c), P7 (Fig. S2d), P9 (Fig. 3e), P10 (Fig. S2e), P11 (Fig. S2f), and P12 (Fig. 3f) are predicted to spontaneously form amyloid fibrils and adopt a predominantly antiparallel alignment within the β-sheet structures in the DMD simulations. Peptides P3 (Fig. 3b), P4 (Fig. S2b), P6 (Fig. S2c), and P8 (Fig. 3d) do not self-assemble to form amyloid-like structures by themselves and remain in a random-coil state at the end of the DMD simulations.

Sequence patterns and signatures of Amyloid-forming peptides

Based on the bioinformatic information gained from the FoldAmyoid predictions and the DMD simulation results, we can make informed predictions as to whether or not a 7-mer peptide is prone to form an amyloid fibril at room temperature. As mentioned earlier, 300 potential amyloid-forming peptides, viz. the top-scoring 150 peptides each from cases 1 to 2, identified via the PepAD algorithm are characterized using the FoldAmyloid webserver to evaluate their intrinsic aggregation propensities. The peptides with at least five residues along the chains above the self-aggregation threshold scale of 21.4 are stored for further study, while the others are abandoned. Approximately 100 *in silico* peptides for case 1 and 140 *in silico* peptides for case 2 meet this criterion.

The 240 promising *in silico* peptides are sorted according to their residue types (Table S1), leading to five distinct sequence patterns of hydrophobic ("H"), charged ("C"), and polar ("P") residues. As shown in Fig. 4, three sequence patterns are identified for case 1 ["CHHHCHP" (pattern 1), "HHCHPHC" (pattern 2) and "CHHHPHC" (pattern 3)], and two sequence patterns are identified for case 2 ["CHHHHHC" (pattern 4) and "HHCHCHH" (pattern 5)]. Subsequently, we characterize the peptide sequences in each pattern and come up with one sequence signature for pattern 1 "CHHHCHP," one sequence signature for pattern 2 "HHCHPHC," two sequence signatures for pattern 3 "CHHHPHC," four sequence signatures for pattern 4 "CHHHHHC," and three sequence signatures for pattern 5 "HHCHCHH." For example, a signature sequence for pattern 1 "CHHHCHP" is *R*-X_1-X_1-X_1-(*E/D*)-X_1-*S*, where X_1 could be any one of the residues *A*, *V*, *L*, and *I*. There are four signature sequences for pattern 4 "CHHHHHC." One of them is (*R/K*)-X_2-X_2-X_2-X_2-X_2-(*E/D*), where X_2 can be any one of the residues *A*, *V*, *L*, *I*, *F*, *Y*, and *W*. The residue options for X_1 and X_2 are identified as a result of their frequent appearance at corresponding sites among the 240 promising *in silico* peptides. By comparing the sequences of the eight *in silico* peptides (P1, P2, P5, P7, P9, P10, P11, and P12) that contain greater than or equal to five self-aggregating residues predicted by the FoldAmyloid tool with the signature sequences, we see that the peptides P1 (*RLLLEAS*), P2 (*RALLSVE*), and P5 (*ALRLSLE*) in case 1 belong to patterns 1, 3, and 2, respectively; the peptides P7 (*RLLAALE*), P9 (*RWLAVAE*) and P11(*RALLALE*) in case 2 belong to pattern 4, and the peptides P10 (*IARLELA*) and P12 (*ALRLELA*) in case 2 belong to pattern 5. The peptide Aβ (16–22) (*KLVFFAE*), which is well known in both experiments and simulations to form a fibril-like structure at room temperature in solution, conforms to the signature sequence (*R/K*)-X_2-X_2-X_2-X_2-X_2-(*E/D*) of pattern 4.

Experimental measurements on peptides predicted to aggregate *in silico*

On the basis of the outcomes of the DMD/PRIME20 simulations and the FoldAmyloid server, eight peptides (P1, P2, P5, P7, P9, P10, P11, and P12) were purchased and characterized experimentally. The peptides were purchased with acetylated and amidated N- and C-termini, respectively, to avoid terminal charges. Note that terminal charges are not considered in the PRIME20 force field. We performed ThT fluorescence spectroscopy, FTIR, negative-stain TEM, and CD (see the Supplementary Material for details). Experimental results are reported in Fig. 5, Fig. S3, and Fig. S4. ThT fluorescence is a commonly used method of detecting amyloid fibril formation: the ThT dye exhibits increased fluorescence upon binding to β-sheets (53–55). At peptide concentrations of 1 mM and near stoichiometric ThT, we observed increased ThT fluorescence intensity for all peptides except for P9 (Fig. 5a, with data from replicates in Fig. S3). FTIR measurements, conducted at the higher peptide concentration of 10 mM, provided direct evidence of antiparallel β-sheet formation for all peptides (Fig. 5b). Depending on the location of the peaks within the amide I (1700–1600 cm^{-1}) region, it is possible to distinguish between antiparallel and parallel β-sheet structures. Antiparallel β-sheets tend to have a maximum at around 1625 cm^{-1} and a smaller peak at around 1695 cm^{-1}, and these features were observed for all the peptides we tested. In contrast, parallel β-sheets would have displayed only one maximum peak at 1640 cm^{-1} and no peak near 1695 cm^{-1} (56–61). It may appear contradictory that the P9 peptide exhibited the FTIR spectrum of an antiparallel β-sheet despite exhibiting no increase in ThT fluorescence, but the peptide concentration differed between these measurements. To further corroborate amyloid fibril formation, we imaged peptide solutions after 1 hour of assembly at 1 mM peptide concentration using negative stain (uranyl acetate) TEM (Fig. 5d and Fig. S4). TEM imaging revealed amyloid fibrils in all peptide solutions tested, although the peptides that exhibited the lowest levels of ThT fluorescence, P5 and P9, appeared to contain fewer fibrils. Furthermore, TEM images of the P9 reveal a low abundance of short fibrils. Finally, we collected CD spectra at the relatively low concentration of 0.2 mM, and these spectra further highlight the concentration-dependent amyloid formation for these peptides. The CD spectrum of 0.2 mM P12 peptide indicates β-strand secondary structure typical of amyloid, with a single minimum at 220 nm. The CD spectra of other peptide peptides appear consistent with partial contributions from molecules in β-strand and random coil conformations, suggesting that self-assembly was not complete for all peptides tested other than P12 at 0.2 mM and 96 hour of assembly time.

To summarize: ThT fluorescence, FTIR, TEM, and CD measurements provide clear evidence of self-assembly into antiparallel β-sheet fibrils for all of the peptides tested. The computational and experimental outcomes for each peptide are summarized in Table S2. Experimental techniques also vary in their compatibility with different peptide concentrations, and results from peptide solutions prepared at concentrations ranging between 0.2 and 10 mM also suggest that assembly depends on concentration in this range. Among the peptides tested experimentally, the most assembly prone peptide appears to be P12, and the least assembly prone peptides appear to be peptides P5 and P9.

Discussion and Conclusion

The PepAD algorithm has been developed in our lab to design peptide supramolecular architectures. Since Sawaya et al. proposed eight possible classes of cross-β spines consisting of a pair of β-sheet layers, researchers have been enthusiastically searching for amyloid-forming peptides in nature to self-assemble into these cross-β spines (22). The β-sheet structure that was used in this work as an initial structure for our discovery of amyloid-forming peptides (Fig. 2b) is the steric zipper of the 8th class (22); it contains two β-sheet layers, antiparallel-oriented β-strands in each layer, and parallel-oriented β-sheets between the two layers. Thus far, only one 6-mer peptide Aβ (35–40) (*MVGGVV*) and one 7-mer peptide Aβ (16–22) (*KLVFFAE*) have been identified in the literature as forming amyloid fibrils of the 8th class in the experiment. By using the PepAD algorithm, we computationally identified a library of 7-mer amyloid-forming peptides that could potentially assemble into the cross-β spine of the 8th class. The PepAD algorithm provides opportunities to discover peptides with the potential to form the cross-β spines of other classes, including but not

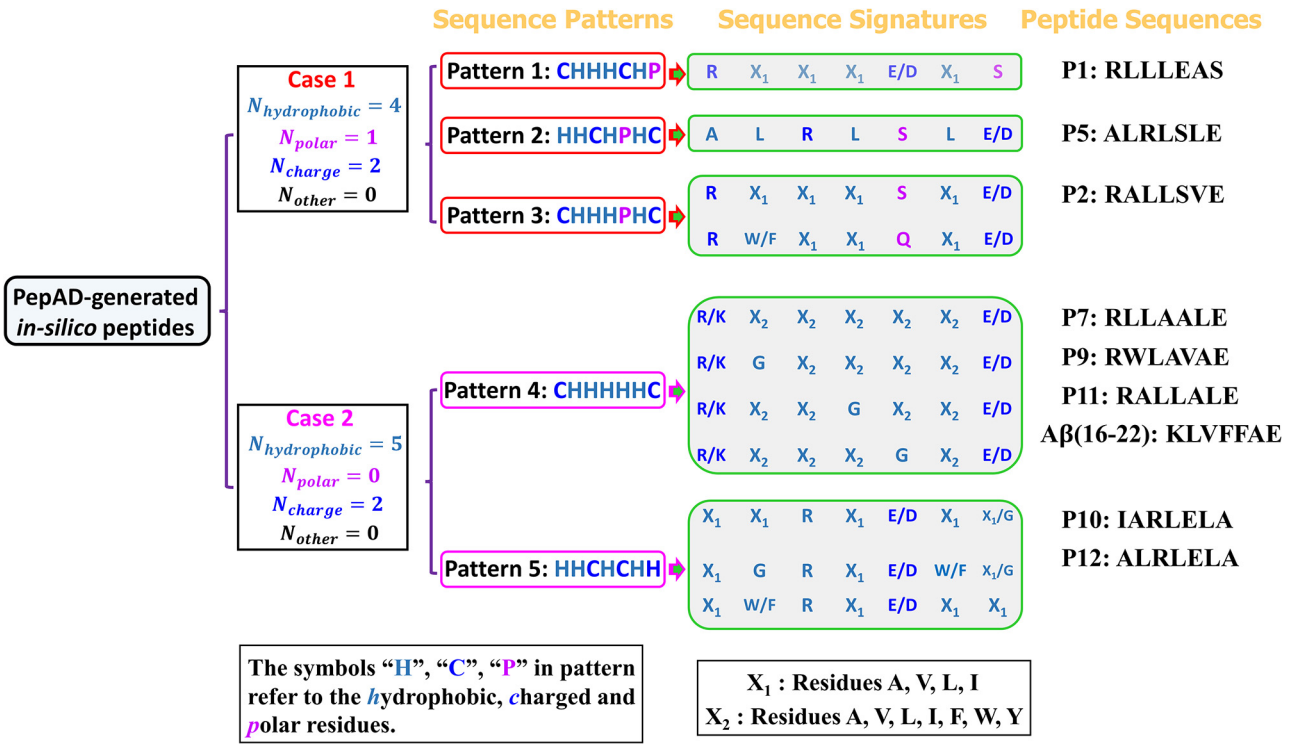

Fig. 4. Sequence patterns and signatures of amyloid-forming peptides, as well as their associated peptides.

limited to the 8th class. In the future, we plan to design peptides that can assemble into other classes.

Our computational studies are devoted to discovering 7-mer amyloid-forming peptides in a fast and efficient manner. By studying the amyloid propensities of twelve PepAD-generated *insilico* peptides, we observed that 7-mer *in silico* peptides with at least five self-aggregation-prone residues, as predicted by the FoldAmyloid tool, aggregate into fibril structures in the DMD/PRIME20 simulations. In contrast, 7-mer peptides having less than five self-aggregation prone residues according to FoldAmyloid do not aggregate into fibril structures in DMD/PRIME20 simulations. Based on the above bioinformatic information, we re-examined each candidate sequence in the sequence pool returned by PepAD and identified approximately 240 potential peptide designs. By sorting out the 240 *in silico* peptide sequences, we obtained five essential sequence patterns as well as associated sequence signatures that seem to suggest which 7-mer peptides will form amyloid. These sequence patterns and signatures suggest a simple initial-stage screen to ascertain computationally whether or not a 7-mer peptide can self-assemble into amyloid fibrils. The present experimental results suggest that computational prediction of peptide assembly can be highly predictive of experimental outcomes.

We have demonstrated that combining the PepAD algorithm, DMD/PRIME20 simulations, and the FoldAmyloid tool makes it possible to generate an experimentally tractable set of 7-mer peptides that can be tested experimentally to discover previously unknown antiparallel β-sheet-forming peptides. Specifically, biophysical measurements revealed that peptides P1, P2, P5, P7, P9, P10, P11, and P12 formed antiparallel β-sheets. The success rate for experimental amyloid-formation of computationally predicted amyloid-forming 7-mer peptides appears to be 100% (eight of eight peptides tested experimentally assembled), although the degree of assembly depended on concentration between 0.2 and 10 mM for all peptides other than P12. Although we are aware of no systematic experimental assessment of assembly for arbitrary 7-amino acid peptides, we suggest that our success rate may be substantially higher than the fraction of all possible 7-mer peptides that would assemble into β-sheets under the conditions tested. However, drawing such a conclusion would require the experimental analysis of a significantly larger set of peptides. Another way to evaluate the PepAD algorithm would be in competition with amino acid sequences chosen by humans experienced in β-sheet peptide design. Since human designs primarily consider simple hydrophobic/polar sequence patterns, one could also compare PedAD peptide designs to randomly selected amino acid sequences with the same hydrophobic/polar patterning. Finally, our experimental results provide a metric for evaluating the PepAD algorithm's design capability and suggest a basis for assessment of future algorithm improvements.

Methods

PepAD algorithm

The PepAD algorithm is a MC-based search procedure to discover peptides that have the potential to self-aggregate into supramolecular architectures, e.g. β-sheet fibril, β-barrel oligomer, or α-helix bundle. A flowsheet of the PepAD algorithm is provided in Fig. S1(a). The procedure is described briefly below.

(1) A predetermined peptide backbone scaffold (herein chosen to be a 2-layer β-sheet fibril) is specified and an initial sequence $S^{(0)}$ is generated on it. In this work, eight 7-mer Aβ (16–22) peptides are aligned in an antiparallel arrangement

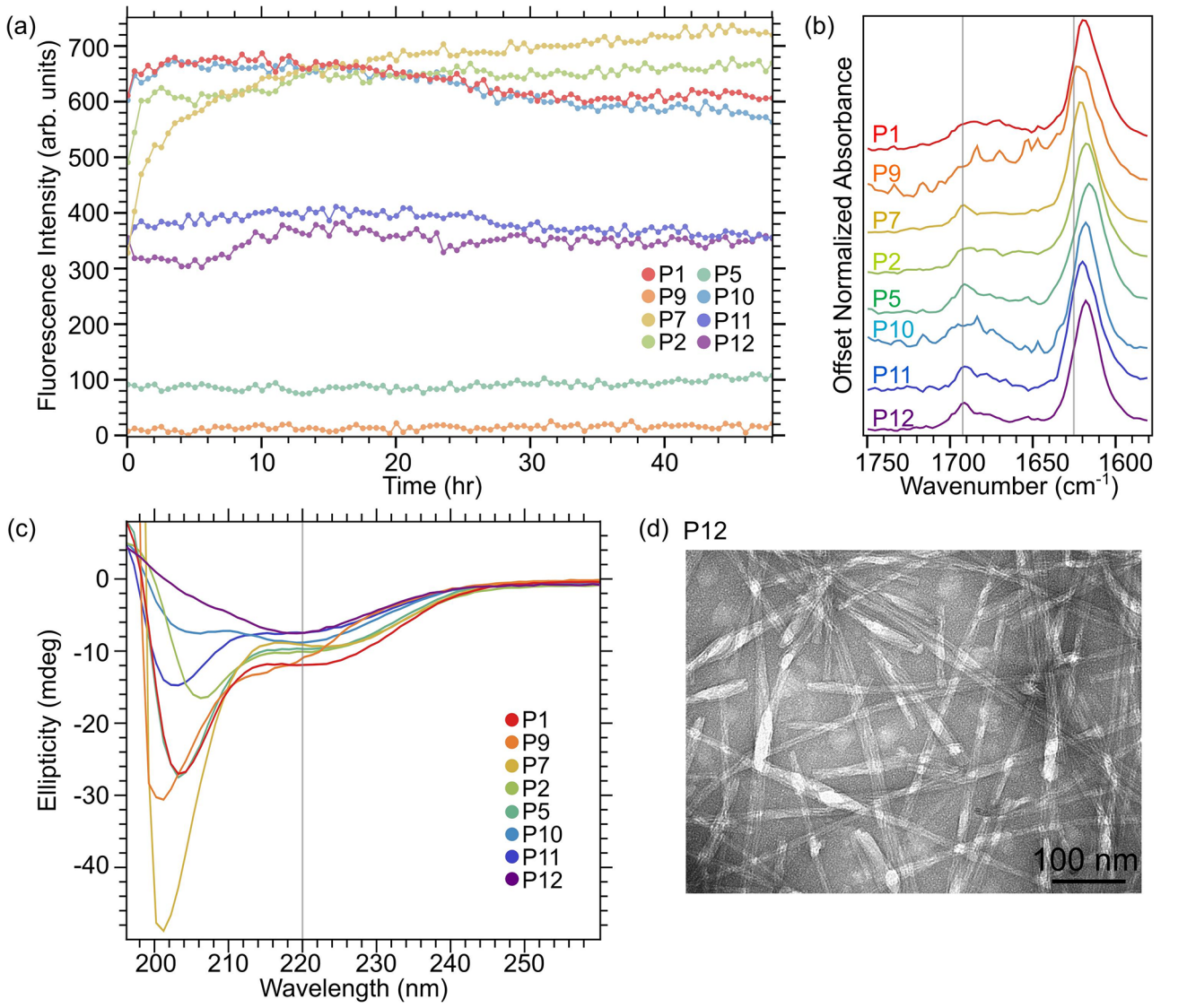

Fig. 5. Experiments on peptides predicted to assemble *in silico*. (a) ThT fluorescence measurements of peptide solutions at 1 mM over 48 hours. Each curve is the average of the replicas, plotted in Fig. S3. (b) FTIR spectra of peptides measured at 10 mM. Gray vertical lines identify peak positions corresponding to antiparallel β-sheets (~1695 and ~1625 cm^{-1}). (c) CD curves measured at 0.2 mM. (d) A TEM image of P12 fibrils (scale bar 100 nm).

within each β-sheet layer (four peptides per layer) and in a parallel arrangement between the β-sheet layers.

(2) The score $\Gamma^{(0)}_{score}$ for this initial sequence $S^{(0)}$ is evaluated using a score function that takes into account the binding affinities between the neighboring chains on the peptide backbone scaffold, as well as the intrinsic aggregation propensities of the individual peptides. The score function is described briefly in "Computational discovery of amyloid-forming peptides" and has been discussed at length in the Supplementary Material.

(3) Two different kinds of trial moves, *viz.* residue mutation and residue exchange, are randomly selected at each evolution step, i, to generate a new trial peptide sequence $S^{(i)}$. A description of the two moves is illustrated in Fig. S1b.

(4) The initial side-chain conformations of the amino acids are chosen from the rotamer library of Lovell et al. (62). Energy minimization is performed to optimize the side-chain configurations of the residues along the peptide chains in the structure. The energy minimization technique is described in our previous work (31).

(5) The score $\Gamma^{(i)}_{score}$ for the new trial sequence $S^{(i)}$ is evaluated, and the new peptide sequence is accepted or rejected using the MC Metropolis sampling method.

(6) Repeat (3, 4, and 5) at least 10,000 times. The sequences with low scores have the potential to self-assemble into the selected supramolecular architecture, e.g., the 2-layer β-sheet fibril.

Two kinds of sequence moves, *viz.* residue mutation and residue exchange, are applied in the PepAD algorithm to perturb the sequences, resulting in new trial peptides. The peptide backbone scaffold is kept fixed during the sequence evolution. The two kinds of sequence moves that are performed are: (i) residue mutation, in which an old residue on all of the peptide chains is randomly chosen and replaced by a new one of the same residue type (hydrophobic, polar, charge, and other); (ii) residue exchange, in which two residues on all of the peptide chains are randomly chosen and swapped, regardless of their residue type. An illustration of the sequence change moves is provided in Fig. S1b. The score function, Γ_{score}, is an essential part of the PepAD algorithm. It is de-

scribed in detail in the Supplementary Material. The 20 standard amino acids have been classified into four distinct residue types in this work. Details are provided in the Supplementary Material.

DMD simulation and PRIME20 model

DMD simulation and the PRIME20 model have been used extensively to examine the kinetics of the fibrillization process for peptides. DMD simulation adopts discontinuous potentials (e.g., hard-sphere and square-well interactions) to describe the interactions between particles, thus allowing for sampling of longer time scales and larger space scales than traditional molecular dynamics simulations. The PRIME20 model is an implicit-solvent coarse-grained protein force field developed in our group that is tailored to simulations of peptide aggregation with DMD. In the PRIME20 model, each amino acid is represented by three backbone spheres (NH–, C_αH–, and CO–) and one side chain sphere (R-). Each side chain of the 20 natural amino acids is assigned a unique size, atomic mass, and C_α-R bond length. Details of the DMD simulations and PRIME20 model are described in our earlier work (47–52).

In this work, we conduct DMD/PRIME20 simulations to examine the aggregation behavior of Aβ (16–22) and the PepAD-generated *in silico* peptides (P1-P12) at $T = 296.1$ K. For each *in silico* peptide system, 48 peptides are placed into a cubic box with edge lengths of 200.0 Å, making the peptide concentration ∼10 mM. Each peptide system starts from a random-coil state and is simulated three times. Each of the DMD simulations is carried out for approximately 5 μs in the canonical ensemble. The Andersen thermostat is implemented to maintain the simulation system at the desired temperature. The fibril cluster of peptides is determined by using the VMD 1.9.3 software to analyze the final simulated structures.

FoldAmyloid Web server

The FoldAmyloid web-server (http://bioinfo.protres.ru/fold-amyloid/) (46) was developed in the Galzitskaya lab to predict which regions on a protein sequence are amyloidogenic. FoldAmyloid is a bioinformatics method that uses statistical data on natural amyloid proteins to obtain inherent aggregation scales for standard amino acids. Residues with predicted self-aggregation scales higher than an empirical threshold value of 21.4 are considered aggregation prone. Details are described in the Supplementary Material.

Peptide Synthesis, FTIR, and ThT

The experimental details of peptide synthesis, ThT fluorescence, FTIR, TEM, CD spectroscopy is described in the Supplementary Material.

Associated Content

Supplementary Material

The Supplementary Material is available free of charge at Detailed descriptions regarding the calculations of score (Γ_{score}), as well as binding free energy ($\Delta G_{binding}$) and intrinsic aggregation propensity ($P_{aggregation}$) in the computational peptide algorithm. The classifications of the 20 standard amino acids into four residue types. The self-aggregation propensities and simulated self-assembled structures of the PepAD-generated *in silico* peptides P2, P4, P6, P7, P10, and P11. The FoldAmyloid-webserver, experimental methods and additional experimental figures are described.

Supplementary Material

Supplementary Material is available at *PNAS Nexus* online.

Authors' Contributions

X.X., A.S.R., A.K.P., and C.K.H. designed research, X.X., A.S.R., S.S, J.V.L., M.E.H., M.J.L., R.G.F., and J.A.E. performed research, X.X., A.S.R., and S.S. analyzed data, X.X., A.S.R., S.S., A.K.P., and C.K.H. wrote the paper.

Funding

The authors in this work would like to thank the National Science Foundation (Award No. OAC-1931430) for financial support. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), supported by National Science Foundation grant number ACI-1548562. X.X. and C.K.H. thank the San Diego Supercomputer Center (SDSC) for providing computing time. The authors also acknowledge the use of instruments in the Materials Characterization Facility and Biopolymer Characterization Core at the Georgia Institute of Technology, as well as the Robert P. Apkarian Integrated Electron Microscopy Core (IEMC) at Emory University. The data described here were collected on the JEOL JEM-1400, 120 kV TEM supported by the National Institutes of Health Grant S10 RR025679, and on the JEOL JEM-2200FS, 200 kV TEM supported by the National Science Foundation Major Research Instrumentation Grant 0923395.

Data Availability

Data are available in the manuscript and/or supporting information. The PepAD code, the top 150 peptide sequences from cases 1 and 2 (300 sequences in total), and the PDB files of peptides P1-P12 obtained from the PepAD algorithm are available at: https://github.com/CarolHall-NCSU-CBE/Data-and-code-for-antiparallel-Self-assembly-design-peptides. All experimental data are available at: http://hdl.handle.net/1853/69976.

References

1. Nagy-Smith K, Moore E, Schneider J, Tycko RT. 2015. Molecular structure of monomorphic peptide fibrils within a kinetically trapped hydrogel network. Proc Natl Acad Sci USA. 112:9816–9816.
2. Dai B, *et al.* 2015. Tunable assembly of amyloid-forming peptides into nanosheets as a retrovirus carrier. Proc Natl Acad Sci USA. 112:2996–3001.
3. Hamley I. 2014. W. Peptide Nanotubes. Angew Chem Int Ed. 53:6866–6881.
4. Branco MC, Pochan DJ, Wagner NJ, Schneider JP. 2010. The effect of protein structure on their controlled release from an injectable peptide hydrogel. Biomaterials. 31:9527–9534.
5. Matson JB, Stupp SI. 2011. Drug release from hydrazine-containing peptide amphiphiles. Chem Commun. 47:7962–7964.
6. Zhao X, *et al.* 2010. Molecular self-assembly and applications of designer peptide amphiphiles. Chem Soc Rev. 39:3480–3498.
7. King PJS, *et al.* 2016. A modular self-assembly approach to functionalized β-sheet peptide hydrogel biomaterials. Soft Matter. 12:1915–1923.
8. Zhang S, *et al.* 1995. Self-complementary oligopeptide matrices support mammalian cell attachment. Biomaterials. 16:1385–1393.

9. Ardejani MS, Orner BP. 2013. Obey the Peptide Assembly Rules. Science. 340:561–562.
10. Hauser CAE, Maurer-Stroh S, Martins IC. 2014. Amyloid-based nanosensors and nanodevices. Chem Soc Rev. 43:5326–5345.
11. Raymod DM, Nilsson BL. 2018. Multicomponent peptide assemblies. Chem Soc Rev. 47:3659–3720.
12. Hartgerink JD, Granja JR, Milligan RA, Ghadiri MR. 1996. Self-assembling peptide nanotubes. J Am Chem Soc. 118: 43–50.
13. Hartgerink JD, Beniash E, Stupp SI. 2001. Self-assembly and mineralization of peptide-amphiphile nanofibers. Science. 294:1684–1688.
14. Frederix PWJM, *et al.* 2015. Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. Nat Chem. 7:30–37.
15. Knowles TPJ, Vendruscolo M, Dobson CM. 2014. The amyloid state and its association with protein misfolding diseases. Nat Rev Mol Cell Biol. 15:384–396.
16. Iadanza MG, Jackson MP, Hewitt EW, Ranson NA, Radford SE. 2018. A new era for understanding amyloid structures and disease. Nat Rev Mol Cell Biol. 19:755–773.
17. Goedert M. 1993. Tau protein and the neurofibrillary pathology of Alzheimer's disease. Trends Neurosci. 16:460–465.
18. Bates G. 2003. Huntingtin aggregation and toxicity in Huntington's disease. Lancet North Am Ed. 361:1642–1644.
19. Li S, *et al.* 2016. Design of asymmetric peptide bilayer membranes. J Am Chem Soc 138:3579–3586.
20. Wilson CJ, *et al.* 2018. Biomolecular assemblies: moving from observation to predictive design. Chem Rev. 118: 11519–11574.
21. Cormier AR, Pang X, Zimmerman MI, Zhou H, Paravastu AK. 2013. Molecular structure of RADA16-I designer self-assembling peptide nanofibers. ACS Nano. 7:7562–7572.
22. Sawaya MR, *et al.* 2007. Atomic structures of amyloid cross-β spines reveal varied steric zippers. Nature. 447:453–457.
23. Tartaglia GG, *et al.* 2008. Prediction of aggregation-prone regions in structured proteins. J Mol Biol. 380:425–436.
24. Oliveberg M. 2010. Waltz, an exciting new move in amyloid prediction. Nat Methods. 7:187–188.
25. Walsh I, Seno F, Tosatto SCE, Trovato A. 2014. PASTA 2.0: an improved server for protein aggregation prediction. Nucleic Acids Res 42:W301–W307.
26. Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L. 2004. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat Biotechnol. 22:1302–1306
27. Rousseau F, Schymkowitz J, Serrano L. 2006. Protein aggregation and amyloidosis: confusion of the kinds? Curr Opin Struct Biol. 16:118–126.
28. Xiao X, Hall CK, Agris PF. 2014. The design of a peptide sequence to inhibit HIV replication: a search algorithm combining Monte Carlo and self-consistent mean field techniques. J Biomol Struct Dyn. 32:1523–1536.
29. Xiao X, Agris PF, Hall CK. 2015. Designing peptide sequences in flexible chain conformations to bind RNA using Monte Carlo, self-consistent mean field and concerted rotation techniques. J Chem Theory Comput. 11:740–752.
30. Xiao X, Agris PF, Hall CK. 2016. Introducing folding stability into the score function for computational design of RNA-binding peptides boosts the probability of success. Proteins. 84: 700–711.
31. Xiao X, Hung ME, Leonard JN, Hall CK. 2016. Adding energy minimization strategy to peptide-design algorithm enables better search for RNA-binding peptides: redesigned λ N peptide binds boxB RNA. J Comput Chem. 37:2423–2435.
32. Xiao X, Wang Y, Leonard JN, Hall CK. 2017. Extended concerted rotation technique enhances the sampling efficiency of computational peptide-design algorithm. J Chem Theory Comput. 13:5709–5720.
33. Spears JL, Xiao X, Hall CK, Agris PF. 2014. Amino acid signature enables proteins to recognize modified tRNA. Biochemistry. 53:1125–1133.
34. Xiao X, *et al.* 2018. Advancing peptide-based biorecognition elements for biosensors using *in-silico* evolution. ACS Sens. 3:1024–1031.
35. Xiao X, *et al.* 2020. In-silico discovery and validation of neuropeptide-Y-binding peptides for sensors. J Phys Chem B. 124:61–68.
36. Xiao X *et al.* 2022. *In-silico* identification of experimental validation of peptide-based inhibitors targeting *Clostridium difficile* toxin A. ACS Chem Biol 17:118–128.
37. Reese HR, *et al.* 2020. Novel peptide ligands for antibody purification provide superior clearance of host cell protein impurities. J Chromatogr A. 1625:461237–461213.
38. Xiao X, *et al.* 2022. De novo discovery of peptide-based affinity ligands for the Fab fragment of human immunoglobulin G. J Chromatogr A. 1669:462941–462915
39. Sarma S, Herrera SM, Xiao X, Hudalla GA, Hall CK, 2022. Computational design and experimental validation of ACE2-derived peptides as SARS-CoV-2 receptor binding domain inhibitors. J Phys Chem B. 126:8129–8139.
40. Gohlke H, Kiel C, Case DA. 2003. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. J Mol Biol. 330:891–913.
41. Xiao X, Zhao B, Agris PF, Hall CK. 2016. Simulation study of the ability of a computationally designed peptide to recognize target $tRNA^{Lys3}$ and other decoy tRNAs. Protein Sci. 25: 2243–2255.
42. Pawar AP, *et al.* 2005. Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. J Mol Biol 350:379–392.
43. Tartaglia GG, *et al.* 2008. Prediction of aggregation-prone regions in structured proteins. J Mol Biol. 380:425–436.
44. Tartaglia GG, Vendruscolo M. 2008. The Zyggregator method for predicting protein aggregation propensities. Chem Soc Rev. 37:1395–1401.
45. Wang Y, *et al.* 2019. Thermodynamic phase diagram of amyloid-β (16-22) peptide. Proc Natl Acad Sci USA. 116: 2091–2096.
46. Garbuzynskiy SO, Lobanov MY, Galzitskaya OV. 2010. FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. Bioinformatics. 26:326–332.
47. Cheon M, Chang I, Hall CK. 2010. Extending the PRIME model for protein aggregation to all 20 amino acids. Proteins. 78:2950–2960.
48. Wang Y, Shao Q, Hall CK. 2016. N-terminal prion protein peptides (PrP (120-144)) form parallel in-register β-sheets via multiple nucleation-dependent pathways. J Biol Chem. 291:22093–22105
49. Nguyen HD, Hall CK. 2004. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. Proc Natl Acad Sci USA. 101:16180–16185.
50. Wagoner VA, Cheon M, Chang I, Hall CK. 2012. Fibrillization propensity for short designed hexapeptides predicted by computer simulation. J Mol Biol. 416:598–609.

51. Wang Y, Latshaw DC, Hall CK. 2017. Aggregation of Aβ(17-36) in the presence of naturally occurring phenolic inhibitors using coarse-grained simulations. J Mol Biol. 429: 3893–3908.
52. Wang Y, *et al.* 2018. Simulations and experiments delineate amyloid fibrilization by peptides derived from glaucoma-associated myocilin. J Phys Chem B. 122:5845–5850.
53. Krebs MRH, Bromley EHC, Donald AM. 2005. The binding of thioflavin-T to amyloid fibrils: localization and implications. J Struct Biol. 149:30–37.
54. Levine H, III. 1993. Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: detection of amyloid aggregation in solution. Protein Sci. 2:404–410.
55. Naiki H, Higuchi K, Hosokawa M, Takeda T. 1989. Fluorometric determination of amyloid fibrils *in vitro* using the fluorescent dye, thioflavine T. Anal Biochem. 177:244–249.
56. Chirgadze YN, Nevskaya NA. 1976. Infrared spectra and resonance interaction of amide-I vibration of the antiparallel-chain pleated sheet. Biopolymers. 15:607–625.
57. Cerf E, *et al.* 2009. Antiparallel β-sheet: a signature structure of the oligomeric amyloid β-peptide. Biochem J. 421:415–423.
58. Sarroukh R, Goormaghtigh E, Ruysschaert J, Raussens V. 2013. ATR-FTIR: a "rejuvenated" tool to investigate amyloid proteins. Biochim Biophy Acta–Biomembr. 1828:2328–2338.
59. Goormaghtigh E, Cabiaux V, Ruysschaert JM. 1994. Determination of soluble and membrane protein structure by fourier transform infrared spectroscopy. I. assignments and model compounds. Subcell Biochem. 23: 329–363.
60. Chan JCC, Oyler NA, Yau W, Tycko R. 2005. Parallel β-Sheets and polar zippers in Amyloid Fibrils formed by residues 10-39 of the yeast prion protein Ure2p. Biochemistry. 44:10669–10680.
61. Cannon MJ, Williams AD, Wetzel R, Myszka DG. 2004. Kinetic analysis of beta-amyloid fibril elongation. Anal Biochem. 328:67–75.
62. Lovell SC, Word M, Richardson JS, Richardson DC. 2000. The penultimate rotamer library. Proteins: Struct. Funct. Genet. 40:389–408.