

Assembly of Stefin B into Polymorphic Oligomers Probed by Discrete Molecular Dynamics

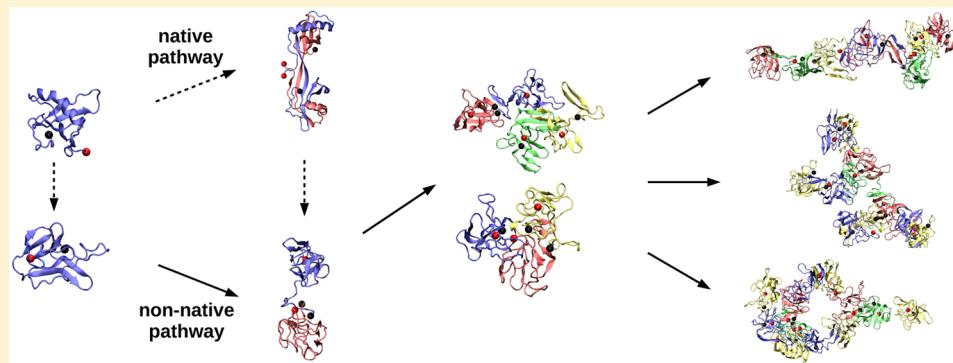
Matjaž Žganec,^{†,‡} Eva Žerovnik,[‡] and Brigita Urbanc^{*,¶,||,†}

[†]Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana, Slovenia

[‡]Department of Biochemistry and Molecular and Structural Biology, Jožef Stefan Institute, 1000 Ljubljana, Slovenia

[¶]Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, United States

 Supporting Information



ABSTRACT: Assembly of an amyloidogenic protein stefin B into molten globule oligomers is studied by efficient discrete molecular dynamics. Consistent with *in vitro* findings, tetramers form primarily through dimer association, resulting in a decreased trimer abundance. Oligomers up to heptamers display elongated rod-like morphologies akin to protofibrils, whereas larger oligomers, decamers through dodecamers, form elongated, branched, as well as annular structures, providing structural insights into pore forming ability and toxicity of amyloidogenic proteins.

INTRODUCTION

The cystatin superfamily consists of structurally homologous proteins that function as cysteine protease inhibitors and are widely distributed in all higher organisms.¹ This superfamily consists of three families: stefins, cystatins, and kininogens.² Human stefins A and B are 98 amino acids-long (11 kDa) globular proteins that lack disulfide bonds and carbohydrates (the IH clan in the MEROPS scheme³). Stefin B acts as a protease inhibitor⁴ and is also involved in lowering oxidative stress,⁵ binding $\text{A}\beta$ both *in vitro* and in cells,⁶ and building the cytoskeleton as a part of a multiprotein complex.⁷ As an intracellular protein, stefin B is present both in the cytoplasm and in the nucleus,⁸ where it binds to histones.⁹ The physiological importance of stefin B is revealed through studies of its naturally occurring mutations, which result in either stefin B deficiency or misfolding, causing a hereditary form of a progressive myoclonal epilepsy of type 1 (EPM1 or Lundborg-Unverricht disease).^{10,11} Stefin B is also present in amyloid plaques together with stefin A, cystatin C, and cathepsins B, H, and L¹² and has been associated with various forms of cancer^{13–15} and autoimmune diseases.¹⁶

Stefins A and B are prone to form amyloid fibrils.⁴ Their native folded structure consists of a five-strand β -sheet wrapped around a single α -helix (N_1 in Figure 9).¹⁷ Under partially denatured *in vitro* conditions, stefins form inactive dimers (N_2

in Figure 9).^{18,19} When fully denatured, stefins A and B form amyloid fibrils.²⁰ Although their sequences are 53% identical, stefin A is significantly more stable than stefin B and forms fibrils only under strongly denaturing conditions (preheating and incubation at $\text{pH} < 2.5$).²⁰ Human stefin B and its variants readily form amyloid fibrils *in vitro*.²¹ Of all the proteins in the cystatin superfamily, the variant of stefin B with amino acid substitutions [C3S, E31Y], stefin B-var2, is the most amyloidogenic under physiologically relevant conditions and has thus been used as an *in vitro* model of amyloid fibril formation.^{20–23} During amyloidogenesis, stefin B-var2 and stefin B variant with a single amino acid substitution [C3S], stefin B-var1, both form on-pathway and off-pathway intermediates.^{24,25} Stefin B-var1 adopts a native as well as a more open folded structure and forms several different dimer conformations.^{25,26} Stefin B-var2 associates into stable domain-swapped dimers²⁷ as well as less stable molten globule-like (MG) oligomers of all sizes.²⁸ Both stefin B-var1 and stefin B-var2 assemble into prefibrillar globular oligomers, ranging in diameter from 5 to 10 nm as shown by size exclusion chromatography (SEC) and dynamic light scattering (DLS).^{27,29} Stefin B-var1, stefin B-var2, and the P74S mutant

Received: January 23, 2015

Published: April 13, 2015

of stefin B-var2 all form MG oligomers, which are cytotoxic.^{27,30} Interestingly, unlike monomers, dimers, and tetramers, which are not toxic per se, the larger oligomers display toxicity in cell cultures,²⁹ indicating a clear relationship between the oligomer order and toxicity. A possible cause of cytotoxicity could be the in vitro observed ability of stefin B variants to form pores in planar lipid bilayers.³¹ The relationship between the on-pathway and off-pathway oligomers formed by stefin B variants is not well understood, and aggregation-prone MG oligomers have not been structurally characterized so far.

Here, a discrete molecular dynamics (DMD) study of stefin B-var2 oligomer formation was conducted to elucidate assembly pathways and structurally characterize oligomers and their morphologies. We employed the implicit-solvent DMD4B-HYDRA force field, which was previously used in studies of folding and oligomerization of two predominant amyloid β -protein ($A\beta$) alloforms, $A\beta40$ and $A\beta42$, relevant to Alzheimer's disease, and their naturally occurring mutants.^{32–37} The strength of effective electrostatic interactions in the DMD4B-HYDRA force field was calibrated to derive the in silico oligomer size distribution consistent with in vitro constraints previously obtained by SEC, DLS, and ionization mass spectrometry.^{25,27} Stefin B-var2 oligomer formation was studied, starting from 8 replicas of 16 unstructured monomers that underwent folding and subsequently assembled into quasi steady-state populations of monomers and oligomers. These stefin B-var2 populations were structurally characterized to provide insights into the structure of oligomers as MG precursors to amyloid fibrils and their corresponding morphologies, which might be in common to a broad class of amyloidogenic proteins and peptides.^{38,39}

■ RESULTS AND DISCUSSION

To examine stefin B-var2 folding and oligomer formation we employ the DMD4B-HYDRA approach, which is based on a four-bead representation of amino acids with three beads representing the amino, α -carbon, and carbonyl groups along the backbone and one bead modeling the entire side chain.⁴⁰ It is well established that protein folding consists of a fast hydrophobic collapse into MG states, followed by a slower conversion from MG states into the native state. Due to a lack of side chain details, the DMD4B-HYDRA approach can capture MG but not native monomer or native assembly conformations. In analogy to folding, protein aggregation also consists of a hydrophobic collapse into MG oligomer assemblies, followed by a slower structural conversion into amyloid fibrils, a process shared by a large number of proteins with no obvious sequence similarities. Overall, neither of these two processes strongly depends on side chain details, thus the DMD4B-HYDRA approach is expected to provide meaningful insights into the structure and dynamics of non-native oligomeric and fibrillar assemblies. Indeed, the DMD4B-HYDRA approach captured essential oligomer size distribution and oligomer structure differences between two 95% identical intrinsically disordered proteins (IDPs), $A\beta40$ and $A\beta42$.^{32,34} Moreover, the resulting DMD4B-HYDRA-derived $A\beta40$ and $A\beta42$ monomers and dimers were structurally favorably compared to their fully atomistic conformations derived by explicit solvent MD.⁴¹ The DMD4B-HYDRA approach provided further biologically relevant insights into the effect of single point mutations^{33,34} and N-terminal truncations³⁶ on $A\beta$ folding and oligomer formation, rendering the approach

suitable for assembly studies of other proteins, such as stefin B-var2.

In vitro results obtained by SEC under amyloidogenic conditions show that in aqueous solutions stefin B-var2 forms native as well as non-native dimers that further assemble into larger oligomers.²⁷ In addition to the abundant dimer population that surpasses the monomer population, the SEC data revealed also a prominent tetramer population and a smaller trimer population.^{25,27} In addition to the relatively stable native dimers, in vitro stefin B-var2 was shown to form transient MG oligomers of different sizes²⁸ that can be examined by the DMD4B-HYDRA approach, in analogy to $A\beta$ assemblies.

The primary structure of the 98-residue-long stefin B-var2 is as follows: MMSGAPSATQ PATAETQHIA DQYRSQLEEK YNKKFPVFKA VSFKSQVVG TNYFIKVHVG DEDFVHLRVL QSLPHENKPL TLSNYQTNKAKHDELTYF.

At neutral pH, stefin B-var2 contains 11 positive and 10 negative charges. The longest purely hydrophobic subsequence is V47–A49. X-ray crystallography revealed the structure of recombinant stefin B monomer in a complex with a cysteine protease papain to be a five-strand β -sheet structure wrapped around the α -helical N-terminal region (P11–V37) and an additional C-terminal strand running along the convex side of the sheet.¹⁷ In our description of the results and discussion, we refer to the three regions of stefin B-var2: the N-terminal region M1–V37 (NTR), the central region F38–L67 (CR), and the C-terminal region R68–F98 (CTR).

Calibration of the Effective Electrostatic Interactions. Unlike the implicit hydrophobicity solvent parameter E_{HP} within the DMD4B-HYDRA approach, which is expected to be similar for all aqueous solutions ($E_{HP} = 0.3$), the other implicit solvent parameter, E_{CH} , which measures the effective electrostatic interaction strength, strongly depends on the solvent, in particular on ionic concentrations used in the experiment of interest, which in turn affect the oligomer size distribution. To adjust E_{CH} to the experimental conditions used in stefin B-var2 studies,^{25,27} DMD simulations were performed at several E_{CH} values. For each value of E_{CH} eight independent trajectories (replicas) of the initial system, consisting of 16 initially unstructured and spatially separated stefin B-var2 peptides, were acquired. In each trajectory, 16 stefin B-var2 peptides underwent folding and subsequent oligomerization until a quasi-steady state was reached. The quasi-steady state of each trajectory was established by monitoring time evolution of the total potential energy and oligomer size distribution. Concomitant with the potential energy and oligomer size distribution, other quantities, including the secondary, tertiary, and quaternary structure, also reached the quasi-steady state. All oligomer size distributions were examined to find the E_{CH} , at which the DMD4B-HYDRA oligomer size distribution was the most consistent with experimental constraints, including the dominance of dimers and tetramers over the trimer population.^{25,27}

Figure 1A demonstrates that the DMD-derived oligomer size distributions shift toward larger oligomer sizes with increasing E_{CH} , and at $E_{CH} \geq 0.5$ the system converges into a single aggregate comprising all 16 peptides (data not shown). A similar tendency to form larger oligomers at increasing E_{CH} has been previously observed in $A\beta$ oligomer formation studies.^{34,42} At $E_{CH} = 0$, the oligomer size distribution was dominated by monomers, dimers, trimers, and only small numbers of larger oligomers (Figure 1A, black histogram), whereas at $0 < E_{CH} \leq$

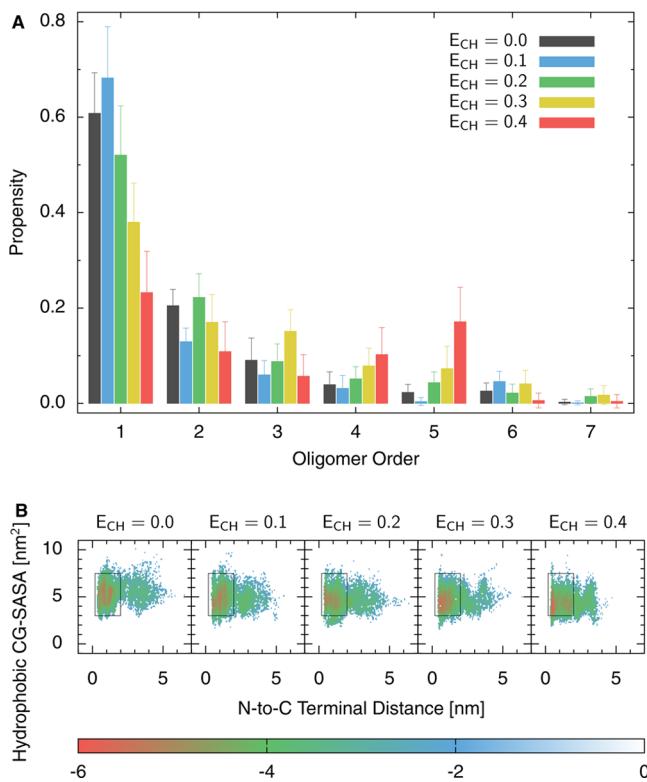


Figure 1. A: Oligomer size distribution of stefin B-var2 at various E_{CH} values. The error bars correspond to the SEM values. B: Free energy landscapes of stefin B-var2 in the space of two reaction coordinates: the distance between the N- and C-terminal C_α beads and the sum of the CG-SASA values over all hydrophobic amino acids. The color scale at the bottom is expressed in $k_B T$. The black rectangular region with the majority of conformations at $E_{CH} = 0$ is plotted to show the effect of increasing E_{CH} on the PMF landscape.

0.4, the number of higher-order oligomers significantly increased, mostly at the expense of monomers and dimers. Among the eight oligomer size distributions corresponding to simulations obtained at the eight different E_{CH} values, the oligomer size distribution at $E_{CH} = 0.4$ was in the best qualitative agreement with the experimental constraints.^{25,27} The in silico dimer abundance was lower than the in vitro dimer abundance, which was expected as the DMD4B-HYDRA-derived stefin B-var2 oligomers contained only non-native dimers. Nonetheless, the dimer and tetramer propensities were both larger than the trimer propensity (Figure 1A, red histogram), consistent with experimental findings.^{25,27}

Characterization of the Oligomer Size Distribution.

Figure 2 shows the time evolution of stefin B-var2 oligomer formation at $E_{CH} = 0.4$. Initially, all peptides were monomeric as reflected in the monomer propensity of 1 at $t = 0$ (Figure 2, black histogram). With time, monomer propensity significantly decreased, whereas oligomer propensities, particularly those of tetramers and pentamers, increased. The oligomer size distribution at $E_{CH} = 0.4$ converged at between 50×10^6 and 60×10^6 simulation time units, and the time frames at $\geq 60 \times 10^6$ time units were used for further structural analysis. When comparing the in silico and in vitro oligomer size distributions, it is important to note that SEC could capture only low molecular weight (LMW) oligomers (up to tetramers) due to limitations of SEC resolution, which sharply decreases with oligomer order. In addition to the LMW oligomers, our

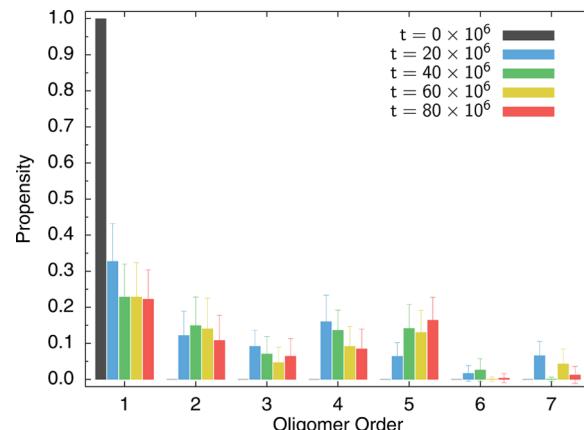


Figure 2. Time evolution of the oligomer size distribution of stefin B-var2 at $E_{CH} = 0.4$. The oligomer size distributions at simulation times: $t = 20 \times 10^6$, $t = 40 \times 10^6$, $t = 60 \times 10^6$, and $t = 80 \times 10^6$ time units were calculated by first averaging the oligomer size distributions within each trajectory (using populations at $t - 5 \times 10^6$, $t - 4.9 \times 10^6$, ..., $t - 0.1 \times 10^6$ time units), then the resulting oligomer size distributions were averaged over the 8 trajectories. The error bars correspond to the SEM values.

simulations revealed also abundant pentamers, nonamers, decamers, undecamers, and pentadecamers (Table 1). These in silico findings are consistent with the more recent report, in which the presence of larger stefin B-var2 oligomers was revealed by ionization mass spectrometry.²⁵

Table 1. Total Number of Conformations at $E_{CH} = 0.4$ Acquired at Time Frames Corresponding to 60.0×10^6 , 60.1×10^6 , ..., 79.9×10^6 Time Units, Used in the Structural Analysis

oligomer order	no. of conformations
1	1080
2	506
3	265
4	476
5	796
6	29
7	21
8	2
9	185
10	357
11	603
12	82
13	1
14	18
15	209
16	15

The Effect of Electrostatic Interaction on the Free Energy Landscape. We examined the effect of E_{CH} on the free energy landscape by calculating the potential of the mean force (PMF) using two reaction coordinates: (i) the distance between the N-terminal and C-terminal C_α atoms and (ii) the sum of coarse-grained solvent-accessible surface area (CG-SASA) values over all hydrophobic amino acids (Figure 1B). In the range of $0 \leq E_{CH} < 0.4$, PMF landscapes spanned a rather extended area in the phase space of the two reaction coordinates. The largest area of the phase space was populated

by an ensemble obtained at $E_{\text{CH}} = 0.2$, for which the PMF landscape contained the most heterogeneous conformational ensemble, characteristic for IDPs. With increasing E_{CH} , a slight shift of the conformational ensemble toward lower hydrophobic CG-SASA values was observed. For $E_{\text{CH}} \geq 0.4$, the PMF landscape covered a more compact area of the phase space. The compactness of the PMF landscape increased with $E_{\text{CH}} > 0.4$ up to $E_{\text{CH}} = 0.8$ (the highest E_{CH} value explored) with the most compact ensemble of conformations (data not shown). Because $E_{\text{CH}} = 0.4$ best matched the experimental constraints, the remaining structural analysis was performed on conformations obtained at $E_{\text{CH}} = 0.4$.

Decrease of Turn and Increase of β -Strand Propensities upon Oligomerization. Turns, β -strands, and coils dominated in the secondary structure of stefin B-var2 monomer and oligomer conformations. The remaining secondary structure elements, such as 3_{10} -helix, π -helix, and isolated bridges, showed significantly lower average propensities of <11%. A comparison of average secondary structure propensities for each oligomer order is shown in Table 2. Overall, the

Table 2. Average Secondary Structure Propensities and Their SEM Values

oligomer order	turn	β -strand	coil
1	0.49 ± 0.01	0.20 ± 0.01	0.21 ± 0.01
2	0.42 ± 0.02	0.26 ± 0.02	0.23 ± 0.02
3	0.40 ± 0.03	0.21 ± 0.03	0.28 ± 0.03
4	0.42 ± 0.02	0.27 ± 0.02	0.22 ± 0.02
5	0.43 ± 0.02	0.26 ± 0.02	0.22 ± 0.02

turn content significantly decreased upon oligomer formation (by ≈ 2.3 SEM) and showed a decreasing trend with increasing oligomer order, whereas the β -strand content increased and reached the highest value in tetramers (Table 2). Dimers, tetramers, and pentamers had significantly higher β -strand content than monomers and trimers. The average coil content was increased in trimers relative to monomer and the other oligomer conformations with similar average coil propensities (Table 2). These secondary structure tendencies were similar to those previously observed in $\text{A}\beta 40$ and $\text{A}\beta 42$ oligomer formation.³² The average α -helix propensity, which was also low in DMD-derived $\text{A}\beta 40$ and $\text{A}\beta 42$ oligomers,³² was different from zero (~ 0.1) only in dimers within the N-terminal region D21–S25, which partially overlaps with the α -helical region (P11–V37) from the X-ray study of the stefin B/papain complex.¹⁷

To elucidate the secondary structure of specific regions, the average turn, β -strand, and coil propensities per amino acid were calculated for conformers up to pentamers (Figure 3, see also Figure S1 in the Supporting Information). The N-terminal region E29–Y31 was characterized by the largest turn propensity in all conformations. In monomers, dimers, and to a lesser degree in larger oligomers, this turn region was flanked by two β -strand regions, L27–E29 and N32–K34. Elevated turn propensities were noted in the CR at K39–Q46, G50–N52, and H58–D61 (Figure S1), and the CTR displayed elevated turn propensities at F70–E76 (the most pronounced in dimers), Q86–N88, and K91–D93. Tetramers showed elevated β -strand propensities at A12–A14 and I19–V23, whereas in pentamers β -strand propensities were elevated at S7–T9 and A12–A14 but not at I19–V23, indicating that distinct N-terminal regions are involved in β -strand formation

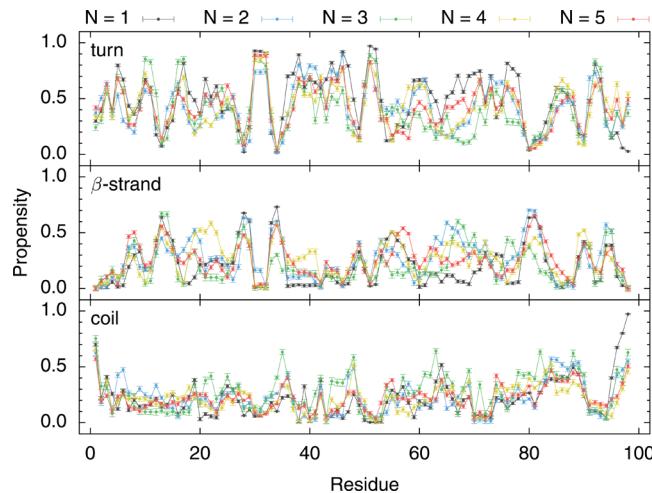


Figure 3. Turn, β -strand, and coil propensities per amino acid of stefin B-var2 for monomers and low-order oligomers obtained from the simulations at $E_{\text{CH}} = 0.4$. The error bars correspond to the SEM values.

in oligomers of different sizes. In addition, the NTR region K33–F35 and the CR region Y53–V57 displayed elevated β -strand propensities in dimers, tetramers, and pentamers but not in trimers. These two β -strand-forming regions, K33–F35 and Y53–V57, may thus be important in stabilizing dimers, tetramers, and pentamers but not trimers, which are less prone to form as evident by their lower abundance in the oligomer size distribution. Notably, these two regions are adjacent to the experimentally characterized linker region (V47–G50), which binds two [P79S]stefin B-var2 monomers into a native domain-swapped dimer.²⁷ Monomers, dimers, and pentamers were characterized by an elevated β -strand content in the CTR at L80–L82. As expected, the least structured regions with the highest coil propensities were the N-terminal M1 and the C-terminal region T96–F98. Coil propensities in the CTR at S83–N88 were high in monomers, dimers, and trimers but significantly decreased in tetramers and pentamers, indicating that the CTR region S83–N88 adopted some structure in tetramers and pentamers. This gain of structure at the C-terminus is consistent with the quenched hydrogen exchange NMR study of stefin B-var1, which reported an increased C-terminal protection from the solvent upon fibril formation.⁴³

Solvent Accessibility Decreases with the Oligomer Order. To gain insights into solvent accessibility of stefin B-var2 monomer and oligomer conformations, CG-SASA was calculated for monomers and oligomers up to pentamers. The average CG-SASA values, which were calculated by averaging the CG-SASA per amino acid over the stefin B-var2 sequence, for different oligomer sizes are summarized in Table 3. As

Table 3. Average CG-SASA per Residue [nm^2] at $E_{\text{CH}} = 0.4$ in Monomer through Pentamer Conformations

oligomer order	CG-SASA [nm^2]
1	0.455 ± 0.002
2	0.408 ± 0.004
3	0.402 ± 0.008
4	0.411 ± 0.004
5	0.396 ± 0.003

expected, the average CG-SASA sharply decreased upon dimer formation. All oligomers had significantly lower solvent exposure than monomers. Pentamers displayed the highest degree of protection from the solvent. To gain more insights into the peptide regions that were the most shielded from the solvent, the average CG-SASA per amino acid was also calculated (Figure 4). Upon oligomer formation, the CG-

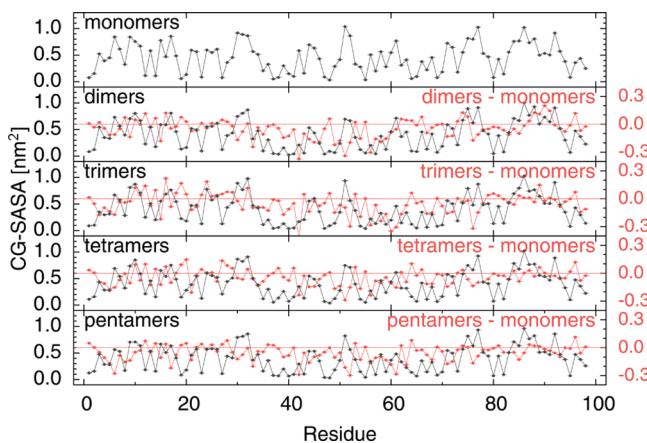


Figure 4. Average CG-SASA (black data set, left y-axis) per amino acid of stefin B-var2 for monomers through pentamers, obtained in simulations at $E_{CH} = 0.4$ and the difference between CG-SASA values in dimer through pentamers with respect to monomers (red data set, right y-axis). The error bars correspond to the SEM values.

SASA values significantly decreased, mostly in the central region of the peptide between F35 and F70, suggesting that this central region was the most involved in oligomer formation. Note that this central region mostly coincides with in vitro characterized interpeptide interface in domain-swapped dimers that becomes solvent inaccessible upon stefin B-var1 dimer formation and remains shielded from the solvent upon further assembly into fibrils.⁴³

Loss of Tertiary Structure upon Oligomer Formation.

Intramolecular contact maps contain information about the tertiary structure of monomer and oligomer conformations. As shown in Figures 5 and 6, C_β contact maps (contacts below the diagonal) were significantly more populated in terms of the average number of contacts than C_α contact maps (contacts above the diagonal), reflecting the effective hydrophobic attraction among hydrophobic C_β beads as the dominant effect in stefin B-var2 folding and assembly. Splitting the stefin B-var2 sequence into three regions (NTR, CR, and CTR), the three diagonal blocks, NTR–NTR, CR–CR, and CTR–CTR, are depicted by black squares on each contact map as a visual guide. In the following, we refer to these diagonal blocks as well as to the remaining off-diagonal blocks, NTR–CR, NTR–CTR, and CR–CTR.

The average number (strength) of intramolecular contacts reached the maximum value (excluding neighboring amino acids in the sequence) of ~ 0.6 and was overall larger than the average number (strength) of intermolecular contacts with a maximum value of ~ 0.3 . In monomers and to a lesser degree in dimers, the N-terminal M1 formed tertiary contacts with I19–A20, consistent with decreased solvent exposure (Figure 4) and a concomitant elevated turn propensity in the NTR at M1–A20 (Figure 3). All intramolecular contact maps displayed well-defined contacts with a propensity of more than 0.5 occurring mostly within the diagonal NTR–NTR and CR–CR blocks.

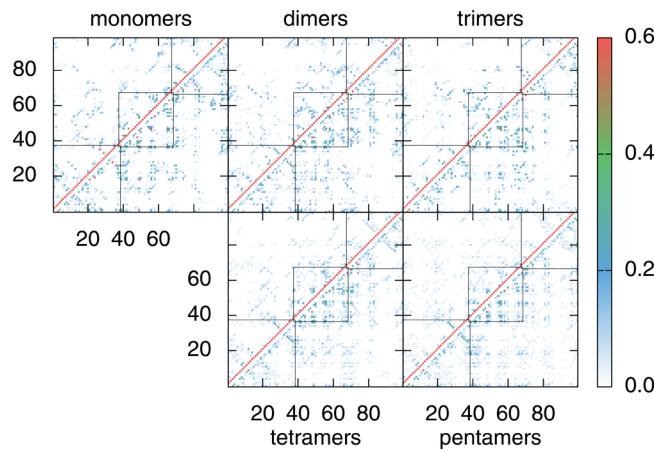


Figure 5. Intramolecular contact maps of stefin B-var2 monomers and low-order oligomers derived from the simulations at $E_{CH} = 0.4$. $C_\alpha–C_\alpha$ contacts are displayed in the upper left triangle, and $C_\beta–C_\beta$ contacts are displayed in the lower right triangle. The color-coded scale is displayed on the right. Three squares along the diagonal correspond to the NTR–NTR, CR–CR, and CTR–CTR blocks.

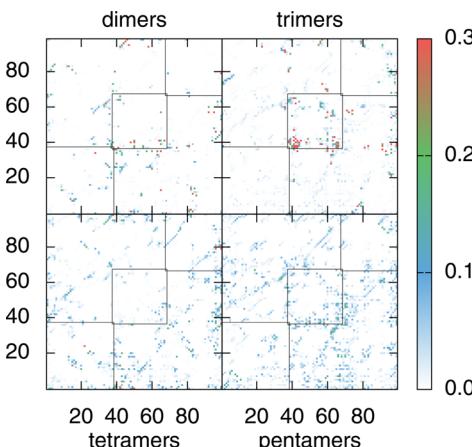


Figure 6. Intermolecular contact maps of low-order oligomers of stefin B-var2 derived from the simulations at $E_{CH} = 0.4$. $C_\alpha–C_\alpha$ contacts are displayed in the upper left triangle, and $C_\beta–C_\beta$ contacts are displayed in the lower right triangle. The color-coded scale is displayed on the right. Three squares along the diagonal correspond to the NTR–NTR, CR–CR, and CTR–CTR blocks.

Apart from these similarities, the tertiary structure of oligomers as determined by intramolecular contact maps was distinct from the tertiary structure of monomers. The number of strong tertiary contacts was by far the largest in monomers, whereas strengths of the most frequent tertiary contacts decreased upon oligomer formation to a maximum value of only ~ 0.2 . Within the diagonal NTR–NTR block, intramolecular contacts formed in monomers with a high propensity, but their strengths (counts) decreased with the oligomer order (Figure 5). Within the diagonal CR–CR block, most intramolecular contacts were preserved across all conformations, although they were significantly stronger in dimers and weaker in pentamers (Figure 5). A tendency of the tertiary structure loss upon oligomer formation was observed in the diagonal CTR–CTR block, where the hairpin structure at F70–L82 in monomers destabilized upon oligomer formation. Within the off-diagonal blocks, in particular in the NTR–CTR, significantly less tertiary contacts were present than within the diagonal blocks. Tertiary

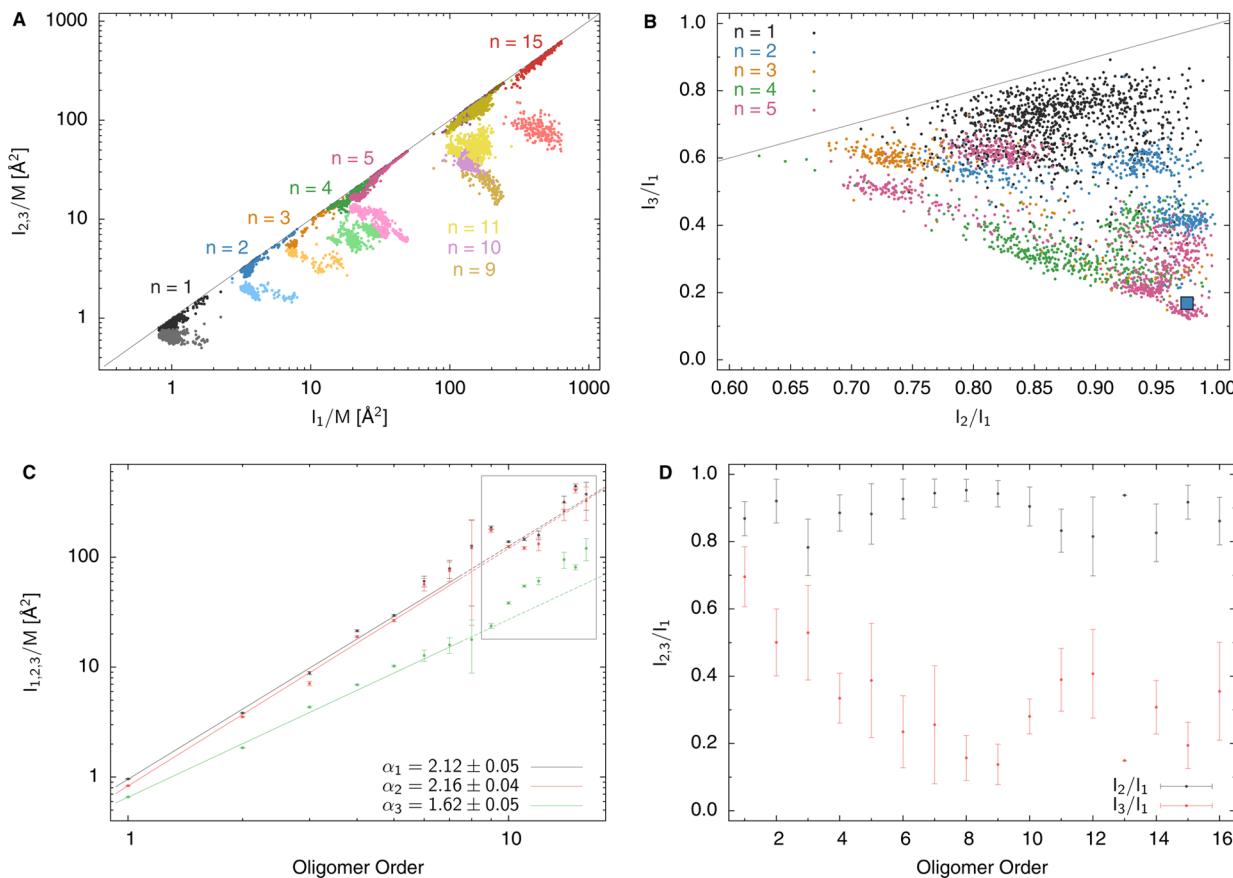


Figure 7. (A) The distribution of the principal moments of inertia for the most populated oligomer orders (black: monomers, blue: dimers, gold: trimers, green: tetramers, magenta: pentamers, brown: nonamers, violet: decamers, yellow: undecamers, red: pentadecamers). The darker color corresponds to I_2 , whereas the lighter color denotes I_3 values. The diagonal line is plotted for reference. The scales on both axes are logarithmic. (B) Two-dimensional distribution of the ratios of the principal moments of inertia relative to the largest moment of inertia (I_1), I_2/I_1 , and I_3/I_1 , for oligomers up to and including pentamers. The diagonal line is plotted for reference. The blue square corresponds to the native stefin A dimer (PDB ID: 1N9J). (C) A log–log plot of the principal moments of inertia versus the oligomer order. A straight line fit included oligomers up to and including heptamers. (D) The two ratios of the principal moments of inertia, I_2/I_1 and I_3/I_1 , versus the oligomer order.

contacts in the off-diagonal blocks that were present in monomers, dimers, and to a lesser extent in trimers were nearly obliterated in tetramers and pentamers. For example, the intramolecular hairpin-like contacts between regions N32–V37 and D93–F98 observed in monomers and dimers were diminished in larger oligomers. The loss of tertiary structure upon oligomer formation, in particular in the off-diagonal blocks, demonstrates that stefin B-var2 monomers underwent significant unfolding prior to oligomer formation as would be expected for a globular protein. These results are in contrast to the tertiary structure of monomers formed by $\text{A}\beta 40$ and $\text{A}\beta 42$ within the DMD4B-HYDRA approach, which underwent only minor changes upon oligomer formation.³²

Oligomers Adopt Distinct Quaternary Structures. The quaternary structure of oligomers was examined through intermolecular contact maps in Figure 6. Cystatins form domain-swapped dimers,^{44–46} which is associated with a loss of their inhibitory activity.⁴⁷ DMD4B-HYDRA-derived stefin B-var2 dimers did not possess many quaternary contacts within the diagonal CTR–CTR block; however, strong localized quaternary contacts were found in the diagonal NTR–NTR and CR–CR blocks (Figure 6). Notable quaternary contacts were found in the off-diagonal NTR–CR and CR–CTR blocks and to a lesser extent also in the NTR–CTR block. Significant quaternary contacts in these off-diagonal blocks concomitant

with a loss of tertiary structure within the diagonal CTR–CTR and NTR–NTR blocks suggests some degree of domain-swapping upon formation of MG-like dimers with remnants of native contacts.

Quaternary contacts in tetramers were more similar to quaternary contacts in dimers than to those in trimers or pentamers, indicating that tetramers formed mostly through dimer association, in agreement with tetramer formation via the hand-shaking mechanism.²⁷ Notably, the diagonal NTR–NTR, CR–CR, and CTR–CTR blocks in dimers and tetramers did not display many strong contacts. In dimers, the CR–CR block contained a few contacts that diminished in tetramers, whereas the NTR–NTR block in dimers displayed fewer quaternary contacts than in tetramers. Importantly, the CTR–CTR block, which barely displayed any quaternary contacts in dimers, contained a well-defined parallel strand of quaternary contacts between regions L67–P74 and P79–Q86, indicating that the CTR, which was free of intermolecular contacts in dimers, actively participated in a dimer to tetramer conversion. Tetramers were characterized by a larger number of quaternary contacts within the off-diagonal blocks than dimers, yet none of these contacts was comparable in strength to the strongest contacts in dimers. The diagonal CR–CR block in tetramers displayed the least quaternary contacts, significantly less than in dimers. The quaternary structure within the CTR–CTR block

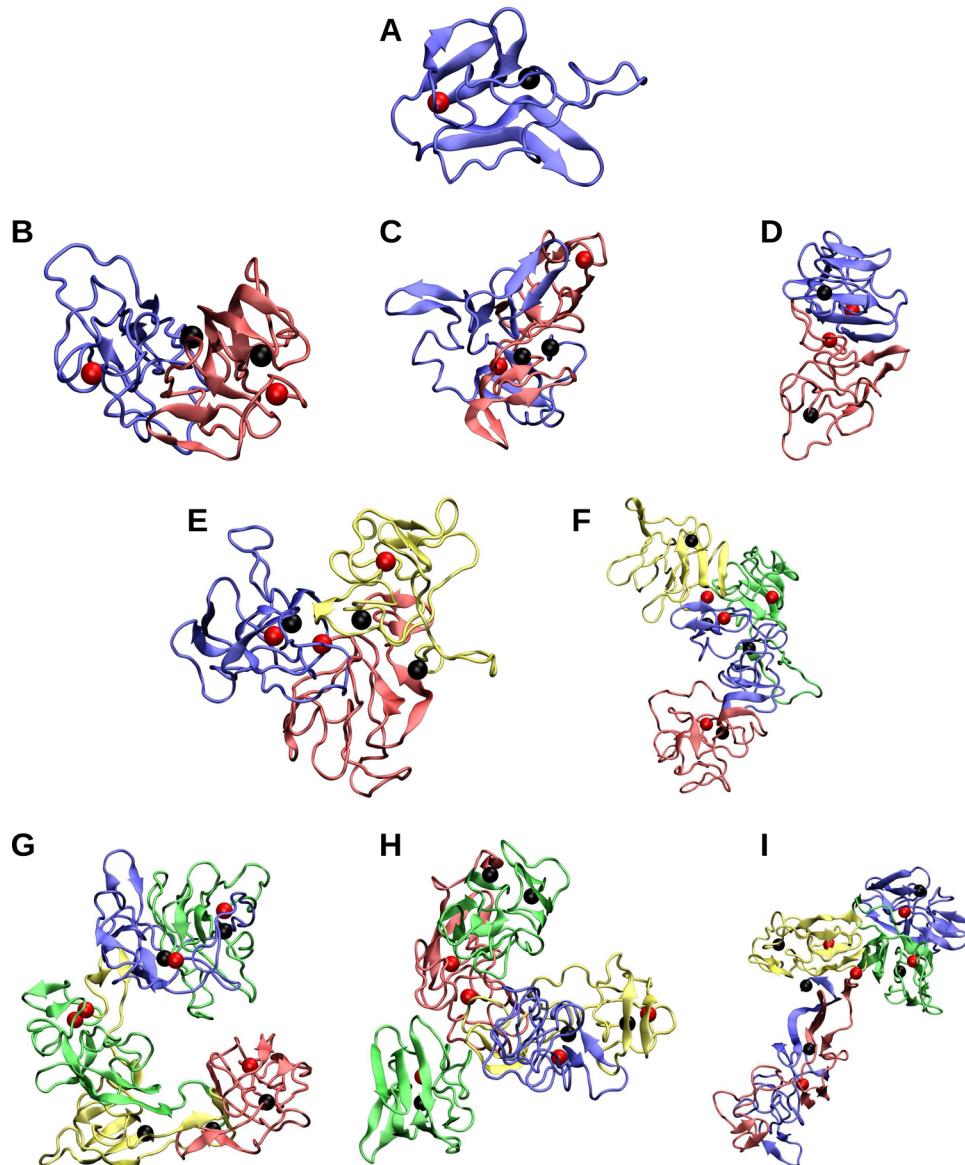


Figure 8. Monomer and oligomer conformations. Characteristic (A) monomer; (B-D) dimer; (E) trimer; (F) tetramer; and (G-I) pentamer conformations with N-terminal and C-terminal C_{α} atoms displayed as red and black spheres, respectively. Dimer conformations included (B) bent, (C) quasi-spherical, and (D) elongated morphologies corresponding to the three blue clusters observed in Figure 7B (top left, top right, bottom right). Pentamer conformations included (G) annular, (H) branched, and (I) elongated morphologies corresponding to the three magenta clusters observed in Figure 7B (left, top middle, bottom right).

(the peptide region L67–Q86) in tetramers that was absent in dimers, concomitant with an overall quaternary structure similarity between dimers and tetramers in regions other than the CTR–CTR block, is reminiscent of the hand-shaking mechanism of dimer to tetramer conversion reported for the P79S mutant of stefin B-var2.²⁷

Quaternary contacts in trimers were quite distinct from those observed in dimers and tetramers. First, the diagonal CR–CR block contained the most prominent contacts, whereas the NTR–NTR block was nearly free of any quaternary contacts. Unlike dimers, trimers displayed quaternary contacts also in the CTR–CTR block. The off-diagonal blocks did not contain as many strong contacts as the CR–CR block but were consistent with an observation that the CR and CTR but not the NTR were involved in trimer formation. Notably, the quaternary structure of trimers that involved the CR and CTR but not the

NTR region was the most consistent with in vitro structural characterization of stefin B-var1 fibrils, in which both the CR and CTR but not the NTR were reported to be protected from the solvent.⁴³ Quaternary contacts in pentamers were unique as they were spread among the diagonal as well as off-diagonal blocks. None of these numerous contacts was particularly strong, and the corresponding quaternary structure resembled simultaneously the quaternary structure of dimers, trimers, and tetramers. Pentamers thus formed through two pathways: (i) a dimer–trimer association and (ii) a monomer-tetramer association, resulting in the most heterogeneous conformational ensemble among the LMW oligomers.

The above analysis showed that stefin B-var2 oligomers possessed distinct quaternary structures, strongly dependent on the oligomer order. Neither of the two characteristics of stefin B-var2 oligomer formation, (i) significant loss of tertiary

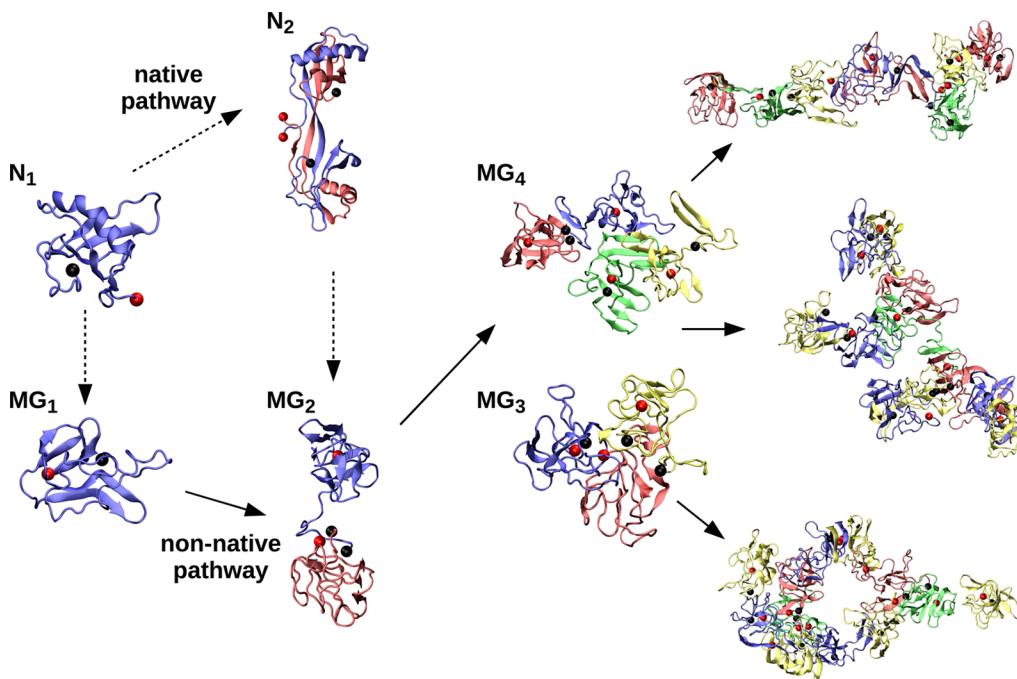


Figure 9. Native (N) and non-native (MG) assembly pathways of stefin B-var2. Under partially denatured conditions, the native monomer (N₁) forms a stable domain-swapped dimer (N₂). Native monomers (N₁) and dimers (N₂) are denatured (dashed arrows) under amyloidogenic conditions, resulting in MG monomer (MG₁) and MG dimer (MG₂) populations that further assemble into MG trimers (MG₃), MG tetramers (MG₄), and larger oligomers of various morphologies (elongated octamer, branched undecamer, and annular dodecamer, from top right to bottom right, respectively).

contacts upon oligomer formation and (ii) the strong dependence of the quaternary structure on the oligomer order, was observed in A β oligomer formation, where the tertiary and quaternary structures remained similar in all LMW oligomers.³²

Characterization of Oligomer Morphology. An overall shape (morphology) of oligomers of a given order was quantified by calculating the average tensor of inertia and finding the three principal moments of inertia $I_1 \geq I_2 \geq I_3$. Only oligomer orders with oligomer populations of more than 100 (within time frames from 60×10^6 – 80×10^6 time units) were included in this calculation to ensure statistical significance (Table 1). A logarithmic plot in Figure 7A reveals characteristic “<”-shaped patterns of distributions of principal moments of inertia for all oligomer sizes. The observed decrease of I_2 and I_3 with I_1 , which was particularly notable in pentamers and nonamers, signified a shift from a globular to an elongated morphology (Figure 7A). Decamers and undecamers displayed notably different distributions of I_2 and I_3 versus I_1 than nonamers, indicating morphologies that were distinctly different from the nonamer morphology.

To elucidate the heterogeneity of oligomer morphologies, we characterized each conformation by two ratios I_2/I_1 and I_3/I_1 and constructed the corresponding two-dimensional distributions of monomers and oligomers up to including pentamers (Figure 7B). Monomers (Figure 8A) displayed the most spherical-like morphologies as reflected by the monomer population being the closest to the rightmost point on the diagonal line ($I_2/I_1 = I_3/I_1 = 1$), where all three moments of inertia are equal (Figure 7B, black points). Overall, the degree of elongation increased with the oligomer order, which is reflected in a decrease in the ratio $\min(I_3/I_1)$ with the oligomer order. Conformations on the right side far from the diagonal

line ($I_3 < I_2 \approx I_1$) corresponded to the most elongated, rod-like morphologies. Conformations close to the diagonal line on the left ($I_2 \approx I_3 < I_1$) corresponded to curvilinear morphologies, including highly bent and annular (ring-like) structures. Dimers were found in three distinct clusters (Figure 7B, blue points), corresponding to bent, quasi-spherical, and elongated morphologies (Figure 8B-D). The ratios I_2/I_1 and I_3/I_1 were calculated also for the experimentally characterized domain-swapped dimer of stefin A with a strongly elongated morphology (Figure 7B, blue square at the bottom right). In trimers, a single rather localized cluster of morphologies was observed (Figure 7B, orange points) with a morphology resembling a three-leaf clover (Figure 8E). Tetramers displayed a single yet dispersed cluster (Figure 7B, green points), corresponding to tetramer morphologies resembling two loosely bound dimers (Figure 8F). Pentamers showed the most heterogeneous population of morphologies, falling into three distinct clusters (Figure 7B, magenta points): corresponding to bent (Figure 8G), quasi-spherical (Figure 8H), and strongly elongated morphologies (Figure 8I), respectively.

To get an insight into the time evolution of oligomer morphologies, we examined scaling of I_1 , I_2 , and I_3 with the oligomer order. If self-assembly proceeded through a spherical oligomer growth, I_1 , I_2 , and I_3 should coincide for all oligomer orders and exhibit similar scaling behavior. A disk-like assembly growth would be revealed through I_1 that scales differently from I_2 and I_3 , $I_1 > I_2 \sim I_3$, whereas a rod-like assembly would exhibit similar scaling of I_1 and I_2 and distinctly different scaling of $I_3 < I_1 \sim I_2$. Figure 7C shows the scaling of I_1 , I_2 , and I_3 with the oligomer order n , $I_i \sim n^{\alpha_i}$ ($i = 1, 2, 3$). Fitting the data up to and including heptamers resulted in $\alpha_1 = 2.12 \pm 0.05$, $\alpha_2 = 2.16 \pm 0.04$, and $\alpha_3 = 1.62 \pm 0.05$, indicating that a rod-like growth was the best approximate model of the oligomer growth. Modeling

an oligomer as a rod of length L and thickness D ($L \gg D$), I_1 and I_2 scale with the rod length L , $I_{1,2} \sim nL^2$, and I_3 scales with the rod thickness D , $I_3 \sim nD^2$. Considering that the mass of an oligomer is proportional to the oligomer order, $m \sim n$, the exponents α_i from the fit to our simulation data can be used to find how the length L and thickness D of the rod-like oligomers scale with the oligomer order n , $L \sim n^\alpha$ and $D \sim n^\beta$. The comparison of the exponents resulted in $1 + 2\alpha = (\alpha_1 + \alpha_2)/2$ and $1 + 2\beta = \alpha_3$, thus $\alpha = 0.57$ and $\beta = 0.31$. As α and β are related to the elongation and lateral growth of rod-like oligomers, our result $\alpha > \beta$ is consistent with an observation of protofibril formation of oligomers smaller than octamers with the oligomer length increasing with the oligomer order significantly faster than their thickness. Interestingly, the oligomer volume $V \sim LD^2$ scaled as $V \sim n^{\alpha+2\beta} = m^{1.19 \pm 0.07}$, which is characteristic for a fractal growth. Decamers, undecamers, and dodecamers, however, deviated from the above scaling (Figure 7C, black rectangle). Inspection of the corresponding morphologies showed increased occurrence of annular (ring-like) morphologies, which caused the observed decrease of I_2 and concomitant increase of I_3 for oligomer orders $9 < n < 13$ (Figure 7D).

CONCLUSION

A large number of proteins form amyloid fibrils under amyloidogenic conditions.⁴⁸ Small globular proteins, such as stefins A and B, are distinctly different from IDPs, such as $\text{A}\beta$, α synuclein, amylin, and prion proteins, as they adopt stable native folds and are characterized by a high overall hydrophobicity and low net charge relative to IDPs.⁴⁹ Stefin B-var2 forms stable dimers under partially denatured conditions, assembles into transient MG intermediates, and forms amyloid fibrils under denatured conditions.²⁸ The relationship between MG oligomers and fibrils is still not well understood.²⁵ While IDPs and their conversion into fibrils have been studied extensively, it is less clear how natively folded (globular) proteins convert to amyloid fibrils. Here, we examined stefin B-var2 assembly into MG oligomers by the DMD4B-HYDRA approach,⁴⁰ which was previously applied to study oligomerization of several naturally occurring $\text{A}\beta$ isoforms.^{32,34,36} Using the available in vitro constraints on stefin B-var2 oligomer size distribution,^{25,27,29} we calibrated the DMD4B-HYDRA force field and then examined its oligomer formation pathways and resulting oligomer morphologies. Just like in $\text{A}\beta$ oligomer formation,³⁴ the overall β -strand propensity in LMW stefin B-var2 oligomers increased with the oligomer order. In contrast to the tertiary and quaternary structures in $\text{A}\beta$ oligomers, which did not strongly depend on the oligomer order,³² stefin B-var2 oligomers displayed a loss of the tertiary structure upon oligomerization and oligomer order-specific quaternary structure. Similarities between dimer and tetramer quaternary contacts indicated that tetramer formation was dominated by dimer association, in agreement with in vitro data.²⁷ Scaling of oligomer morphologies with the oligomer order revealed a rod-like growth, whereby oligomers elongated along the long principal axis faster than growing laterally. Pentamers and in particular decamers through dodecamers displayed a high degree of morphological polymorphism, including highly elongated protofibril-like conformations (Figure 9, octamer at the top right), wide, laterally branched structures (Figure 9, undecamer at the middle right), and, importantly, annular (ring-like) conformations (Figure 9, dodecamer at the bottom right), which demonstrate that larger oligomers of stefin B-var2

can adopt pore-like conformations akin to those observed in vitro in planar lipid bilayers.³¹ The ability of oligomers larger than pentamers to form pore-like structures in our simulations correlates with the experimentally observed onset of cytotoxicity of stefin B-var2 hexamers and larger oligomers²⁹ and thus supports the hypothesis that oligomer pore formation by amyloidogenic proteins causes membrane disruption that results in cell death.^{50,51} Because both stable native and aggregation-prone MG dimers of stefin var2-B coexist in vitro, our findings here are consistent with a hypothesis that under amyloidogenic conditions, the native assembly pathway into dimers acts as a kinetic trap but coexists with a non-native pathway that proceeds through formation of MG oligomers, which retain some native contacts (both pathways are shown in Figure 9). Formation of MG oligomers is then followed by a nucleated structural conversion into fibrils as observed within a minimal self-assembly model,⁵² in in vitro $\text{A}\beta$ aggregation studies,^{53,54} and supported by experimental findings on stefin B-var1 fibril formation.⁴³

METHODS

The DMD approach, which combines the four-bead peptide model with backbone hydrogen bonding and implicit-solvent amino acid-specific interactions (DMD4B-HYDRA force field), has been previously described and applied to simulations of $\text{A}\beta$ oligomer formation and folding of mucin domains.^{32,34,40,55}

The DMD4B-HYDRA Approach. *Discrete Molecular Dynamics.* If interactions between particles are approximated with one or more square wells, the fixed time step-driven molecular dynamics (MD) method can be replaced by a computationally more efficient DMD approach.^{56,57} DMD is driven by discrete events—collisions between particles. Between successive collisions each particle moves with constant velocity eliminating the need for explicit numerical integration of the equations of motion. However, an efficient sorting algorithm is required to find the shortest time, which determines the next collision. Our DMD approach implements periodic boundary conditions, constant temperature through Berendsen thermostat,⁵⁸ and uses fixed simulation volume and number of particles.

Four-Bead Protein Model with Backbone Hydrogen Bonding. In the four-bead peptide model, each amino acid is represented by four beads (except for glycine modeled by three backbone beads), which is the minimal representation capturing the chiral nature of each amino acid (except glycine). The three beads for the amino (N), the α -carbon (C_α), and the carbonyl (C) group are used to model the peptide backbone. The fourth bead (C_β) represents the entire side chain of each amino acid except glycine, which lacks the side chain bead. Bond lengths and their angular constraints were determined phenomenologically from known folded structures of ~7,700 proteins obtained from the Protein Data Bank.⁵⁹ Covalent bond lengths and their angular constraints were allowed to fluctuate by 2% about their average values. The effective hydrogen bond interaction between the backbone nitrogen atom N_i of the i -th amino acid residue and the carbon atom C_j of the j -th amino acid residue was implemented to allow for the secondary structure formation and stability.⁶⁰ The strength of the backbone hydrogen bond interaction E_{HB} represents a unit of energy, which sets the temperature unit to E_{HB}/k_B .

The Implicit-Solvent DMD4B-HYDRA Force Field. The implicit-solvent amino acid-specific force field, DMD4B-HYDRA, was implemented into the four-bead peptide model

by introducing the effective amino acid-specific interactions among C_{β} beads due to hydrophobic properties of individual amino acids.³² Using the phenomenological Kyte-Doolittle hydrophobicity scale,⁶¹ in which each amino acid is associated with a hydrophobicity parameter, at neutral pH the C_{β} beads of the amino acids I, V, L, F, C, M, and A are treated as hydrophobic; the C_{β} beads of the amino acids N, Q, and H are treated as noncharged hydrophilic; and the C_{β} beads of the amino acids R, K, D, and E are treated as charged hydrophilic. For the remaining amino acids, the absolute value of the hydrophobicity parameter in the Kyte-Doolittle hydrophobicity scale is below the threshold value, and the corresponding C_{β} beads interact in the DMD4B-HYDRA force field only through the excluded volume interaction. To derive the force field interaction parameters, each C_{β} bead is assigned a normalized hydrophobicity parameter with a value between -1 (the most hydrophobic) and 1 (the most hydrophilic). The effective interaction between pairs of hydrophobic/hydrophilic C_{β} beads is attractive/repulsive with the strength that is equal to an average of the two corresponding hydrophobicity parameters. All other interactions between two C_{β} beads are due to the excluded volume.

The effective hydrophobic interaction between two C_{β} beads is modeled by a single square well potential with the interaction range of 7.5 Å and the depth equal to the strength of effective hydrophobic interactions, E_{HP} , multiplied by the average hydrophobicity parameter of the two interacting C_{β} beads. The strength of effective hydrophobic interactions is set to $E_{HP} = 0.3$, the value which captured the distinct oligomer size distributions of A β 40 and A β 42, the two predominant A β alloforms relevant to Alzheimer's disease, in aqueous solutions.^{32,34} This implementation of the effective hydrophobic interactions in the absence of explicit water molecules ensures that the hydrophobic side chain beads cluster together and thereby minimize their overall CG-SASA, whereas the hydrophilic side chain beads repel each other to maximize their overall CG-SASA, mimicking the presence of water.

The effective electrostatic interaction between two charged C_{β} beads is implemented as a double square-well potential with a maximum depth E_{CH} and two interaction ranges, 6 Å and 7.5 Å, respectively, as described previously.⁴⁰ To compare the predictions of the DMD4B-HYDRA approach to experimental findings, the E_{CH} was calibrated to ionic strength of the aqueous solution. The case $E_{CH} = 0$ describes a solution, in which the charges are completely screened, resulting in no effective electrostatic interactions among the charged C_{β} beads. Note that in this case the repulsive interactions among all hydrophilic C_{β} beads, including the charged ones, are different from zero. In our simulations, we used a wide range of values of the implicit solvent parameter E_{CH} from 0 to 0.8, which allowed us to calibrate our simulation model to the experimental solvent conditions that were used to measure the in vitro oligomer size distribution of stefin B-var2.

Simulation Protocol. Our system was comprised of 16 noninteracting and spatially separated unfolded stefin B-var2 proteins arranged in a simple cubic lattice in a cubic simulation box of side 250 Å. At a given value of the implicit solvent parameter E_{CH} , eight distinct initial populations of unfolded monomers were obtained by performing high-temperature DMD simulations at $T = 4$ and saving conformations every 0.1×10^6 simulation steps. The resulting ensembles of 16 peptides were then used as initial configurations for the constant-volume production runs, which were performed at physiological temperature $T = 0.13$.^{33,34} We acquired 8 distinct 80×10^6

simulation-step-long trajectories of stefin B-var2 assembly at each of the following strengths of the effective electrostatic interaction E_{CH} : 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, resulting in 64 trajectories total. In each trajectory, we monitored the assembly from initially separated and unstructured monomeric peptides into an ensemble of monomers and oligomers of different sizes until the oligomer size distribution reached a quasi-steady state, which occurred between 50×10^6 and 60×10^6 time units. With $T = 0.13$ modeling the physiological temperature (310 K) and $k_B T = 0.6$ kcal/mol, the unit of energy $E_{HB} = 4.6$ kcal/mol. As the spatial resolution in the DMD4B-HYDRA model is 10^{-2} nm, each time unit is equal to ~ 0.3 ps, so that the total simulation time of each trajectory, 80×10^6 time units, corresponds to $24 \mu s$.

Structural Analysis. Ensembles of conformations acquired at time frames ranging from 60×10^6 to 80×10^6 time units were used to feed the structural analysis. Each configuration at a given time frame contained a mixture of monomers and oligomers of various sizes. First, the oligomer size was determined by calculating the distances between centers of atoms that belonged to different chains. If the smallest such distance between two chains was ≤ 5 Å, the two chains were assigned to be within the same oligomer. The same cutoff distance of 5 Å was used for the calculation of intra- and intermolecular contact maps to elucidate the tertiary and quaternary structure of conformers. Second, all configurations were converted into PDB⁵⁹ files containing individual monomers or oligomers. In this process, the backbone oxygen and the backbone amide hydrogen were added, which allowed VMD to identify backbone hydrogen bonds and thus secondary structure elements. Third, the PDB files were grouped by the oligomer order. Our structural analysis was limited to monomers and oligomers up to pentamers, which were abundant enough to obtain statistically reliable results (Table 1). The morphological analysis also included the more abundant larger oligomers: nonamers, decamers, undecamers, and pentadecamers.

Oligomer Size Distribution. For each value of the implicit solvent parameter E_{CH} , we calculated the oligomer size distribution by averaging over the ensembles of conformations between time frames 60.0×10^6 and 79.9×10^6 for each of the 8 trajectories containing 16 stefin B-var2 proteins, resulting in $5 \times 8 \times 16 = 640$ protein chains contributing to each oligomer size distribution. The size distribution was first calculated for each trajectory, and then the averages and standard error of the mean (SEM) values were derived by considering eight independent distributions.

Secondary Structure. The propensities of the secondary structure elements, β -strand, turn, coil, and α -helix, were calculated using the STRIDE^{62,63} program within the Visual Molecular Dynamics (VMD) software package.⁶⁴ This analysis included the time frames at 60.0×10^6 , 60.1×10^6 , ..., 79.9×10^6 simulation steps within each of the eight trajectories corresponding to the implicit solvent parameter $E_{CH} = 0.4$.

Potential of the Mean Force. For each value of the implicit solvent parameter E_{CH} , the PMF calculation included the ensembles of conformations at 60.0×10^6 , 60.1×10^6 , ..., 79.9×10^6 simulation steps within each of the eight trajectories with 16 peptides, that is $200 \times 8 \times 16 = 25,600$ peptides. First, we selected two reaction coordinates to describe the conformation of each peptide: the distance between the N- and C-terminal C_{α} atoms and the sum of CG-SASA values over all hydrophobic residues. N-to-C terminal distance and CG-SASA of hydro-

phobic residues were chosen as two independent phase space variables that measure the end-to-end distance distribution of each peptide and its solvent accessibility. Then we analyzed each time frame and binned all conformers into a two-dimensional array as determined by the two reaction coordinates to create a two-dimensional histogram of populations. Based on the density of conformers in each bin, we selected $0.05 \text{ nm} \times 0.1 \text{ nm}^2$ as the final two-dimensional bin box size. N-to-C terminal distances of conformers ranged up to 6 nm. Almost all conformers had CG-SASA of hydrophobic residues below 10 nm^2 . Finally, we transformed the conformer density histogram to free energy landscape by applying $F_i = -k_B T \ln(N_i/N)$, where N_i is the number of conformers in bin i and $N = \sum N_i$ the total number of conformers in the histogram.

Coarse-Grained Solvent Accessible Surface Area. CG-SASA per amino acid was calculated within VMD⁶⁴ using the time frames at 60.0×10^6 , 60.1×10^6 , ..., 79.9×10^6 simulation steps for each of the eight trajectories corresponding to the implicit solvent parameter $E_{\text{CH}} = 0.4$. Briefly, a spherical surface around each bead is created, followed by the CG-SASA calculation for each amino acid. In the four-bead model, a combined solvent-exposed surface area of all four beads, the backbone oxygen and amide hydrogen of each amino acid that did not overlap with surface areas of any other beads, was then calculated.

Contact Maps. A contact map was defined as a matrix of the average number of contacts between residues of the same protein chain (intramolecular or tertiary contacts) or residues of different protein chains within the oligomer (intermolecular or quaternary contacts). According to our definition, two residues formed a contact if the distance between the corresponding C_α or C_β beads was $\leq 5 \text{ \AA}$. C_α (C_β) contact map values for residues i and j were thus calculated as the average number of $C_{\alpha,i}$ - $C_{\alpha,j}$ ($C_{\beta,i}$ - $C_{\beta,j}$) contacts. We calculated intramolecular contact maps of monomers and both intra- and intermolecular contact maps of dimers, trimers, tetramers, and pentamers. The contact maps described here did not distinguish among the contacts that formed as a result of effective hydrophobic or electrostatic interaction, hydrogen bonding, or a proximity to such contacts. The contact maps were averaged over the eight trajectories acquired for $E_{\text{CH}} = 0.4$ using time frames 60.0×10^6 , 60.1×10^6 , ..., 79.9×10^6 .

Principal Values of the Moment of Inertia Tensor. We calculated the tensor of the moment of inertia as implemented in VMD for all conformations in the populations of monomers and oligomers acquired with 8 trajectories at $E_{\text{CH}} = 0.4$ in time frames between 60.0×10^6 , 60.1×10^6 , ..., 79.9×10^6 time units. By calculating the eigenvalues of the inertia tensor, we obtained the three principal moments of inertia for each oligomer order that contained more than 100 conformations (monomers through pentamers, nonamers, decamers, undecamers, and pentadecamers).

ASSOCIATED CONTENT

Supporting Information

To show the changes in the secondary structure upon oligomer formation, we plotted the differences in turn, β -strand, and coil propensities per amino acid between oligomers of a certain order and monomers. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Phone (0)215 8952726. E-mail: brigita@drexel.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by Slovenian research agency (ARRS), grant J7-4050 (E.Ž.), Slovene human resources development and scholarship fund, grant 11012-37/2014 (M.Ž.), and bilateral grant by ARRS (E.Ž. and B.U.). The doctoral program of M.Ž. is partially funded by the European Union. The access to Extreme Science and Engineering Discovery Environment (XSEDE) supercomputing facilities through grant TG-PHYS100030 (B.U.), supported by the National Science Foundation, is kindly acknowledged. We thank the Slovenian Initiative for National Grid (SLING) for the access to computational resources.

REFERENCES

- (1) Barrett, A. *Trends Biochem. Sci.* **1987**, *12*, 193–196.
- (2) Turk, V.; Bode, W. *FEBS Lett.* **1991**, *285*, 213–219.
- (3) Rawlings, N.; Tolle, D.; Barrett, A. *Nucleic Acids Res.* **2004**, *32*, D160–D164.
- (4) Turk, V.; Stoka, V.; Turk, D. *Front. Biosci.* **2008**, *13*, 5406–5420.
- (5) Lehtinen, M. K.; Tegelberg, S.; Schipper, H.; Su, H.; Zukor, H.; Manninen, O.; Kopra, O.; Joensuu, T.; Hakala, P.; Bonni, A.; Lehesjoki, A.-E. *J. Neurosci.* **2009**, *29*, 5910–5915.
- (6) Škerget, K.; Taler-Verčič, A.; Bavdek, A.; Hodnik, V.; Čeru, S.; Tušek-Žnidarič, M.; Kumm, T.; Pitsi, D.; Pompe-Novak, M.; Palumaa, P.; Soriano, S.; Kopitar-Jerala, N.; Turk, V.; Anderluh, G.; Žerovnik, E. *J. Biol. Chem.* **2010**, *285*, 3201–3210.
- (7) Giamo, R. D.; Riccio, M.; Santi, S.; Galeotti, C.; Amrosetti, D. C.; Melli, M. *Hum. Mol. Genet.* **2002**, *11*, 2941–2950.
- (8) Kopitar-Jerala, N. *Curr. Protein Pept. Sci.* **2012**, *8*, 767–775.
- (9) Čeru, S.; Konjar, Š.; Repnik, U.; Križaj, I.; Benčina, M.; Renko, M.; Nepveu, A.; Žerovnik, E.; Turk, B.; Kopitar-Jerala, N. *J. Biol. Chem.* **2010**, *285*, 10078–10086.
- (10) Laioti, M. D.; Mirotsou, M.; Buresi, C.; Peitsch, M. C.; Rossier, C.; Ouazzani, R.; Baldy-Moulinier, M.; Bottani, A.; Malafosse, A.; Antonarakis, S. E. *Am. J. Hum. Genet.* **1997**, *60*, 342–351.
- (11) Genton, P. *Epilepsia* **2010**, *51*, 37–39.
- (12) Ii, K.; Ito, H.; Kominami, E.; Hirano, A. *Virchows Arch. A: Pathol. Anat. Histopathol.* **1993**, *423*, 185–194.
- (13) Kos, J.; Krašovec, M.; Cimerman, N.; Nielsen, H. J.; Christensen, I. J.; Brünner, N. *Clin. Cancer Res.* **2000**, *6*, 505–511.
- (14) Friedrich, R.; Jung, K.; Türk, I.; Rudolph, B.; Hampel, G.; Schnorr, D.; Loening, S. A. *Eur. J. Cancer* **1999**, *35*, 138–144.
- (15) Zore, I.; Krašovec, M.; Cimerman, N.; Kuhelj, R.; Werle, B.; Nielsen, H. J. *Biol. Chem.* **2001**, *382*, 805–810.
- (16) Cimerman, N.; Meško-Brguljan, P.; Krašovec, M.; Šuškovič, S.; Kos, J. *Clin. Chim. Acta* **2001**, *310*, 113–122.
- (17) Stubbs, M. T.; Laber, B.; Bode, W.; Huber, R.; Jerala, R.; Lenarčič, B.; Turk, V. *EMBO J.* **1990**, *9*, 1939–1947.
- (18) Žerovnik, E.; Jerala, R.; Kroon-Žitko, L.; Turk, V.; Lohner, K. *Eur. J. Biochem.* **1997**, *245*, 364–372.
- (19) Jerala, R.; Žerovnik, E. *J. Mol. Biol.* **1999**, *291*, 1079–1089.
- (20) Jenko, S.; Škarabot, M.; Kenig, M.; Gunčar, G.; Mušević, I.; Turk, D.; Žerovnik, E. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 417–425.
- (21) Žerovnik, E.; Pompe-Novak, M.; Škarabot, M.; Ravnikar, M.; Mušević, I.; Turk, V. *Biochim. Biophys. Acta* **2002**, *1594*, 1–5.
- (22) Žerovnik, E.; Turk, V.; Walther, J. P. *Biochem. Soc. Trans.* **2002**, *30*, 543–547.

- (23) Žerovník, E.; Zavašník-Bergant, V.; Kopitar-Jerala, N.; Pompe-Novak, M.; Škarabot, M.; Goldie, K.; Ravník, M.; Muševič, I.; Turk, V. *Biol. Chem.* **2002**, *383*, 859–863.
- (24) Škerget, K.; Vilfan, A.; Pompe-Novak, M.; Turk, V.; Waltho, J. P.; Turk, D.; Žerovník, E. *Proteins: Struct., Funct., Bioinf.* **2009**, *74*, 425–436.
- (25) Taler-Verčič, A.; Kirsipuu, T.; Friedmann, M.; Noormägi, A.; Polajnar, M.; Smirnova, J.; Žnidarič, M. T.; Žganec, M.; Škarabot, M.; Vilfan, A.; Staniforth, R. A.; Palumaa, P. *Int. J. Mol. Sci.* **2013**, *14*, 18362–18384.
- (26) Paramore, R.; Morgan, G. J.; Davis, P. J.; Sharma, C. A.; Hounslow, A.; Taler-Verčič, A.; Žerovník, E.; Waltho, J. P.; Cliff, M. J.; Staniforth, R. A. *Front. Mol. Neurosci.* **2012**, *5*, 94.
- (27) Kokalj, S. J.; Gunčar, G.; Štern, I.; Morgan, G.; Rabzelj, S.; Kenig, M.; Staniforth, R. A.; Waltho, J. P.; Žerovník, E.; Turk, D. J. *Mol. Biol.* **2007**, *366*, 1569–1579.
- (28) Žerovník, E.; Jerala, R.; Kroon-Žitko, L.; Pain, R. H.; Turk, V. J. *Biol. Chem.* **1992**, *267*, 9041–9046.
- (29) Čeru, S.; Kokalj, S.; Rabzelj, S.; Škarabot, M.; Gutierrez-Aguirre, I.; Kopitar-Jerala, N.; Anderluh, G.; Turk, D.; Turk, V.; Žerovník, E. *Amyloid* **2008**, *15*, 147–159.
- (30) Čeru, S.; Žerovník, E. *FEBS Lett.* **2008**, *582*, 203–209.
- (31) Rabzelj, S.; Viero, G.; Gutierrez-Aguirre, I.; Turk, V.; Serra, M. D.; Anderluh, G.; Žerovník, E. *FEBS J.* **2008**, *275*, 2455–2466.
- (32) Urbanc, B.; Cruz, L.; Yun, S.; Buldyrev, S. V.; Bitan, G.; Teplow, D. B.; Stanley, H. E. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17345–17350.
- (33) Lam, A.; Teplow, D. B.; Stanley, H. E.; Urbanc, B. *J. Am. Chem. Soc.* **2008**, *130*, 17413–17422.
- (34) Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. *J. Am. Chem. Soc.* **2010**, *132*, 4266–4280.
- (35) Urbanc, B.; Betnel, M.; Cruz, L.; Li, H.; Fradinger, E.; Monien, B. H.; Bitan, G. *J. Mol. Biol.* **2011**, *410*, 316–328.
- (36) Meral, D.; Urbanc, B. *J. Mol. Biol.* **2013**, *425*, 2260–2275.
- (37) Attar, A.; Meral, D.; Urbanc, B.; Bitan, G. In *Bio-nanoimaging: Protein Misfolding and Aggregation: Assembly of Amyloid β -Protein Variants Containing Familial Alzheimer's Disease-Linked Amino Acid Substitutions*; Uversky, V., Lyubchenko, Y., Eds.; Elsevier: 2014; Chapter 38, pp 429–442.
- (38) Uversky, V.; Fink, A. *Biochim. Biophys. Acta* **2004**, *1698*, 131–153.
- (39) Gerber, R.; Tahiri-Alaoui, A.; Hore, P. J.; James, W. *J. Biol. Chem.* **2007**, *282*, 6300–6307.
- (40) Urbanc, B.; Borreguero, J. M.; Cruz, L.; Stanley, H. E. *Methods Enzymol.* **2006**, *412*, 314–338.
- (41) Barz, B.; Urbanc, B. *PLoS One* **2012**, *7*, e34345.
- (42) Yun, S.; Urbanc, B.; Cruz, L.; Bitan, G.; Teplow, D. B.; Stanley, H. E. *Biophys. J.* **2007**, *92*, 4064–4077.
- (43) Morgan, G. J.; Giannini, S.; Hounslow, A. M.; Craven, C. J.; Žerovník, E.; Turk, V.; Waltho, J. P.; Staniforth, R. A. *J. Mol. Biol.* **2008**, *375*, 487–498.
- (44) Staniforth, R.; Giannini, S.; Higgins, L.; Conroy, M.; Hounslow, A.; Jerala, R.; Craven, C.; Waltho, J. *EMBO J.* **2001**, *20*, 4774–4781.
- (45) Janowski, R.; Kozak, M.; Jankowska, E.; Grzonka, Z.; Grubb, A.; Abrahamson, M.; Jaskolski, M. *Nat. Struct. Biol.* **2001**, *8*, 316–320.
- (46) Sanders, A.; Craven, J.; Higgins, L. D.; Giannini, S.; Conroy, M. J.; Hounslow, A. M. *J. Mol. Biol.* **2004**, *336*, 165–178.
- (47) Ekiel, I.; Abrahamson, M. *J. Biol. Chem.* **1996**, *271*, 1314–1321.
- (48) Chiti, F.; Dobson, C. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.
- (49) Uversky, V.; Gillespie, J.; Fink, A. *Proteins: Struct., Funct., Genet.* **2000**, *41*, 415–427.
- (50) Lashuel, H.; Hartley, D.; Petre, B.; Walz, T.; Lansbury, P. T., Jr. *Nature* **2002**, *418*, 291.
- (51) Quist, A.; Doudevski, I.; Lin, H.; Azimova, R.; Ng, D.; Frangione, B.; Kagan, B.; Ghiso, J.; Lal, R. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10427–10432.
- (52) Barz, B.; Urbanc, B. *J. Phys. Chem. B* **2014**, *118*, 3761–3770.
- (53) Bitan, G.; Lomakin, A.; Teplow, D. B. *J. Biol. Chem.* **2001**, *276*, 35176–35184.
- (54) Lee, J.; Culyba, E.; Powers, E.; Kelly, J. *Nat. Chem. Biol.* **2011**, *7*, 602–609.
- (55) Barz, B.; Turner, B. S.; Bansil, R.; Urbanc, B. *J. Biol. Phys.* **2012**, *38*, 681–703.
- (56) Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1959**, *31*, 459–466.
- (57) Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1960**, *33*, 1439–1451.
- (58) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *8*, 3684–3690.
- (59) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (60) Ding, F.; Borreguero, J. M.; Buldyrev, S. V.; Stanley, H. E.; Dokholyan, N. V. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 220–228.
- (61) Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105–132.
- (62) Frishman, D.; Argos, P. *Proteins: Struct., Funct., Genet.* **1995**, *23*, 566–579.
- (63) Heinig, M.; Frishman, D. *Nucleic Acids Res.* **2004**, *32*, W500–W502.
- (64) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.