

Examining the Ensembles of Amyloid- β Monomer Variants and Their Propensities to Form Fibers Using an Energy Landscape Visualization Method

Murilo N. Sanches,^{||} Kaitlin Knapp,^{||} Antonio B. Oliveira Jr., Peter G. Wolynes, José N. Onuchic, and Vitor B. P. Leite*



Cite This: *J. Phys. Chem. B* 2022, 126, 93–99



Read Online

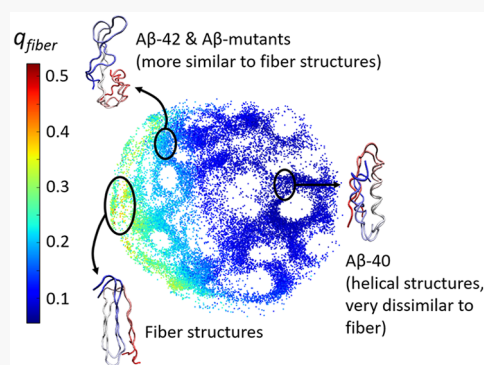
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: The amyloid- β ($A\beta$) monomer, an intrinsically disordered peptide, is produced by the cleavage of the amyloid precursor protein, leading to $A\beta$ -40 and $A\beta$ -42 as major products. These two isoforms generate pathological aggregates, whose accumulation correlates with Alzheimer's disease (AD). Experiments have shown that even though the natural abundance of $A\beta$ -42 is smaller than that for $A\beta$ -40, the $A\beta$ -42 is more aggregation-prone compared to $A\beta$ -40. Moreover, several single-point mutations are associated with early onset forms of AD. This work analyzes coarse-grained associative-memory, water-mediated, structure and energy model (AWSEM) simulations of normal $A\beta$ -40 and $A\beta$ -42 monomers, along with six single-point mutations associated with early onset disease. We analyzed the simulations using the energy landscape visualization method (ELViM), a reaction-coordinate-free approach suited to explore the frustrated energy landscapes of intrinsically disordered proteins. ELViM is shown to distinguish the monomer ensembles of variants that rapidly form fibers from those that do not form fibers as readily. It also delineates the amino acid contacts characterizing each ensemble. The results shed light on the potential of ELViM to probe intrinsically disordered proteins.



INTRODUCTION

Most of molecular biology has been guided by the structure–function paradigm, whereby the three-dimensional structure of a protein determines specific biological functions. The process by which a newly created peptide sequence can achieve an organized structure can be quite complex but is now understood by considering the energy landscape theory.^{1–5} More recently, a broad class of proteins whose amino acid sequences result in only minimally structured proteins has been recognized. Commonly referred to as intrinsically disordered proteins (IDPs),^{6–10} these peptides appear to explore rough, relatively flat energy landscapes that are marked by a multitude of local basins having similar free energies.^{11–14} These proteins lack structure when fully solvated but may undergo large conformational changes to more ordered states under certain conditions.^{15–17} However, until such triggering conditions are met, they explore a large conformational space.^{18–20} Some IDPs while disordered tend to exist in relatively compacted conformations,²¹ which are the result of competition between entropy and energy. Computer simulations have provided insight into the conformational transitions they can undergo. A major problem in such studies is how to analyze and classify the large and highly variable data sets that are produced.²²

A significant number of neurological diseases have been associated with IDPs.^{23–26} Often, IDP monomers can interact

to form aggregate structures, ranging over a large scale of sizes, which either may still be disordered or may be highly ordered. Determining the mechanisms driving the aggregation process from individual unstructured monomers to structured fibrils is critical to understanding the role of aggregates in disease.²⁷ Over the last 20 years, a huge number of research efforts have been directed toward understanding the aggregation processes associated with Alzheimer's disease (AD), a form of dementia suffered by an estimated 50 million people worldwide.²⁸ One of the proteins closely associated with AD is a peptide fragment known as amyloid- β ($A\beta$). The most abundant forms of $A\beta$ are either 40 or 42 residues in length.²⁹ The apparent neural damage in AD has long been thought to be directly caused by the presence of large fibers and plaques,³⁰ but small, soluble oligomers of $A\beta$ ^{30,31} have also been considered the source of the damage. Understanding $A\beta$ in its monomeric form is a useful place to start in any case.^{32–34} Two regions in the monomers have been identified as playing key roles in the

Received: September 28, 2021

Revised: December 9, 2021

Published: December 30, 2021



dynamics of aggregation, the central hydrophobic core (CHC), consisting of residues 17–20, and the B-turn region formed by residues 23–28. The B-turn region is capable of forming a salt-bridge, which has been observed to accelerate the aggregation process.³⁵ Despite differing by only two amino acids in length, A β -40 and A β -42 markedly differ in their aggregation tendencies. While A β -40 is naturally present at a ratio of 10:1 over A β -42 in cerebrospinal fluid, A β -40 is only present in the much smaller ratio of 1:3 in amyloid plaques as compared to A β -42,³⁶ consistent with the experimental observation that the nucleation and elongation rates of A β -42 are much faster than those of A β -40.³⁷ Studies have previously shown that the behavior of these isoforms diverge early in the aggregation process during the formation of small oligomers.³⁸ The aggregation rate for A β -42 correlates with the β -strand content found in its monomeric form.³⁹ It should also be noted that several single-point mutations of amyloid- β have been identified⁴⁰ that are associated with early onset and inherent forms of the disease.^{41–43}

To investigate the correlation of the monomer structural ensembles with the aggregation propensities of A β , we have used the energy landscape visualization method (ELViM)⁴⁴ to analyze the conformational ensembles of A β -40 and A β -42 monomers, as well as the landscapes of six A β -40 mutations. The ELViM technique has already exhibited a capacity to discriminate and clarify the energy landscapes of IDPs, as shown in a previous work with the prostate associated gene 4 (PAGE4).⁴⁵ Through the local structural similarity metric proposed by Wolynes et al.,⁴⁶ it is possible to survey and triangulate a high-dimensional conformational phase space and then, using ELViM, project the ensembles to two optimal dimensions while preserving local proximities. This allows an intuitive visual analysis of the energy landscape.

METHODS

Simulations Details. Choosing a force field to simulate a free monomer of A β , in both its 40 and 42 residue variants, is crucial. The ideal force field for modeling disordered peptides should have a strong physical foundation, incorporate information from ideal geometries and known electrostatic properties, and be shown to make accurate structure predictions. For this reason, we used the associative-memory, water-mediated, structure and energy model (AWSEM)⁴⁷ force field. This coarse-grained force field was obtained by using energy-landscape-based machine learning, run in LAMMPS,³⁹ and can explore a large energy landscape. The A β -42 simulations were initiated starting from a solution structure of A β -42 (1Z0Q) inferred from NMR.⁴⁰ In the case of A β -40, the decision was made to use the same starting structure, though truncated to 40 residues. Comparable solution studies of A β -40 (2LFM)⁴¹ also show significant local helical secondary structure and therefore were started from the same conditions. Each simulation was initialized by energy minimization with a 2 fs time step for the first 100 000 time steps, followed by a simulation run of 10⁶ time steps. During this run, structures were sampled every 1000 time steps. Simulations were carried out at 200, 250, 300, 350, and 400 K, with two independent simulations per specified temperature. This resulted in a total of 0.4 μ s of simulation time and 100 000 sampled structures. For the mutant A β -40 species, simulations were also started from the same annealed structure to avoid any secondary structure bias, since the structures of the monomeric mutants have yet to be resolved.

The simulation procedure was consistent with that previously described, with the data from the full range of temperatures used to determine the energy landscape, though only simulations carried out at 300 K were used for structural sampling and comparison.

Energy Landscape Visualization Method. There are many methods for the analysis of complex multidimensional data, and the two most well-known approaches are principal component analysis (PCA) and multidimensional scaling (MDS) methods. PCA describes multivariate data in terms of effective principal vectors. The input data of each sample or state is a vector in N -dimensional space, and all of the states are projected onto the principal directions, in which the data displays the highest variance. In this low-dimensionality representation, one seeks a meaningful clustering of the data.^{48,49} On the other hand, MDS methods rely on pairwise distances between elements.^{50,51} MDS algorithms are non-linear dimensionality reduction that can be broadly classified into two categories and that seeks to capture global or local properties of the data.⁵² The energy landscape visualization method (ELViM) belongs to the MDS class of technique, and a combination of appropriate metrics and a projection scheme allows obtaining insight into global properties while preserving local neighborhood features.⁵³

For globular folded proteins, conformational changes can usually be represented in terms of an effective reaction coordinate associated with a unique native state. This approach is not as useful for IDPs because their landscapes are not strongly funneled. To visualize the details of the energy landscape of A β , we instead employ the ELViM, which starts with a metric based on the internal distances between pairs of amino acids.^{44,54} One measure of similarity between two conformations k and l that is quite useful is given by

$$q^{k,l} = \frac{1}{N} \sum_i \sum_j \exp \left[-\frac{(r_{ij}^k - r_{ij}^l)^2}{\sigma_{ij}^2} \right] \quad (1)$$

where N is the total number of pairs of residues, r_{ij}^k is the distance between the residue i and j in the conformation k , r_{ij}^l is the distance in the conformation l , and σ_{ij} is a weighing constant equal to $\sigma_0 |i - j|^\epsilon$, with $\sigma_0 = 1$ and $\epsilon = 0.15$. The normalized distance between the conformations can then be defined by $\delta^{k,l} = 1 - q^{k,l}$.

The methodology is divided into three steps: (I) The first step to create a visualization of the landscape is to calculate the distances between every pair of structures in the trajectory, building an $n \times n$ dissimilarity matrix. (II) The second step consists of a data clustering based on the distance between the conformations in the matrix. Structures close to each other by less than a given cutoff distance are clustered into a single conformation. (III) The last step is the multidimensional reduction of the original data to a 2D or 3D projection using these clusters as a guide. A multidimensional projection technique initially places the objects of the matrix at random initial positions in a 2- or 3-dimensional space and then applies a minimization step to preserve the distances in the dissimilarity matrix. The process stops when the difference in distances between consecutive iterations is below a given cutoff. ELViM analyzed data included conformations of A β -40 and A β -42 sampled for 250, 300, and 350 K to ensure sampling the conformation phase space as detailed as possible. However, the ELViM has a limitation on the number of conformations, so we had to restrict the total number of points. Since we

wanted to probe all of the mutants in the same analysis, we restricted the chosen conformations to those obtained at 300 K, which are more meaningful and more significantly weighted. We end up having a less accurate description of mutants compared to the A β -40 and A β -42, with which we used a much larger data set. However, it still indicates the qualitative differences of the mutants when compared to the wild type A β -40 and the A β -42.

WHAM and Weighted Conformations. Structures of each sequence from all temperatures (200, 250, 300, 350, and 400 K) were analyzed using the weighted histogram analysis method (WHAM). By an iteration procedure and using data sampled under different temperatures, WHAM estimates the density of states $\Omega(E)$ that satisfies the expected probability for the data obtained in each condition.⁵⁵ 20K conformations for each temperature were used to estimate $\Omega(E)$. The probability of a state with energy i at inverse temperature β is given by

$$P(E_i, \beta) = \frac{\Omega(E_i)e^{-\beta E_i}}{\Omega_T} \quad (2)$$

where

$$\Omega_T = \sum_i \Omega(E_i) \quad (3)$$

We then divided the states of each ensemble into bins: in 1D for the case of the q_{fiber} coordinate, or in the ELViM projection for a given 2D grid. We consider all of the m states within a given bin b , and the free energy of b will be given by

$$F^b = -\frac{1}{\beta} \ln \left(\sum_{i=1}^m \frac{\Omega(E_i)e^{-\beta E_i}}{\Omega_T} \right) \quad (4)$$

RESULTS AND DISCUSSION

Since A β is an intrinsically disordered protein, describing conformational changes would require more than one reaction coordinate. ELViM is a reaction-coordinate-free method that allows one to visualize in an optimal way the high-dimensional conformational phase space of each type of A β in a small number of dimensions. By analyzing all of the sampled conformations at once, we can make a comparison of the conformational ensembles of different studied monomers. We create a large projection with the A β -40, A β -42, and A β mutant conformations, all represented and analyzed in a single phase space. Figure 1 shows this projection with the different species separated in different colors in panel A and the entire projection labeled in colors indicating the radius of gyration in panel B. Each conformation of the data set is represented by a point mapped in the 2D projection. The entire data set with conformations from different systems is analyzed together, at once. The axes are the same in all views. Through this procedure, we can analyze each system data set separately and make a direct comparison between them. The axes have no particular meaning; the important aspect is the relative distance between conformations. Conformations are not evenly distributed in the 2D projection, which is expected since it reflects the molecular dynamics conditions in which the data is sampled, favoring some regions of the phase space according to their energies. Moreover, the dimensionality of the system is so high that some sampled conformations may be considered “outliers” compared with the bulk of conformations of a local region. Such outliers end up being considerably

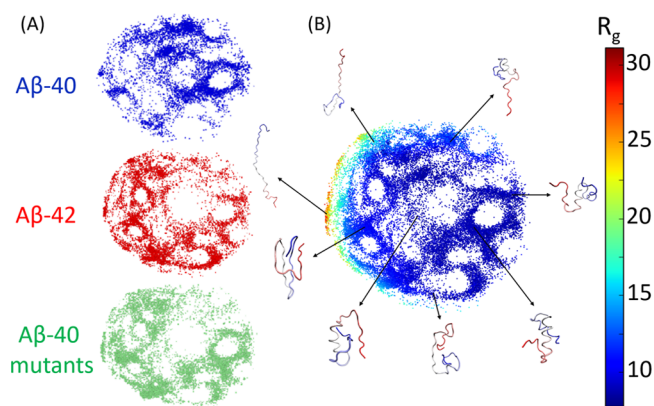


Figure 1. Conformational phase space of the simulated structures visualized through ELViM. (A) Distinct species present in the projection, with the A β -40 in blue, A β -42 in red, and A β -40 mutants in green. (B) Each conformation as a function of the radius of gyration, with the most significant conformations of each region displayed around it.

distant from the other local conformations, and they appear as “islands” surrounded by void areas in the 2D projection, as can be seen in Figure 1B. An alternative analysis could remove such outliers from the data set, but that would represent an unnecessary filter in the data set. In our analysis, we have avoided such data filtering. One can see that most of the A β -40 points are located on the right side of the projection, while the A β -42 points are more widely spaced, and the mutants follow a similar behavior. The structures around the projection in Figure 1B exemplify the conformations of each region, which shows that the regions most populated by A β -40 correspond to the α -helix clusters, which are also more compact. In contrast, the regions with a bigger radius of gyration correspond to the A β -42 and A β mutants.

This result is supported by Figure 2, which shows the ELViM projection as a function of the q coordinate measured with respect to the A β -40 monomer (2LMN) in the fiber state.

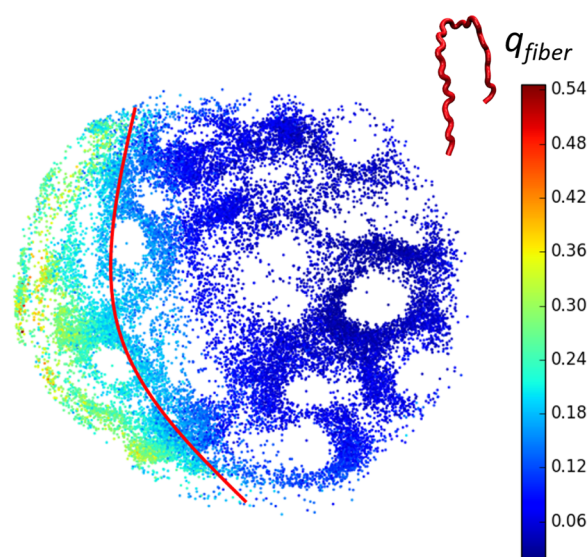


Figure 2. Two-dimensional ELViM projection as a function of q with the fiber monomer as reference (red structure, PDB 2LMN). The red line acts as a divider between a high and low q_{fiber} region in the projection.

We call this variable q_{fiber} , and for any conformation i , using eq 1, $q_{\text{fiber}} = q_{i,\text{fiber}}$, where fiber corresponds to the 2LMN fiber conformation. We see that the $A\beta$ -40 landscape populates a region with low q_{fiber} , while $A\beta$ -42 and the mutant landscapes populate a region with a somewhat high q_{fiber} . This suggests a reason for the much higher propensity for fiber formation by $A\beta$ -42 and the mutants.

To examine this finding more closely, using the weighted histogram analysis method (WHAM), we calculated the free energy as a function of the similarity to the fiber monomer structure. The results shown in Figure 3 suggest that there is an

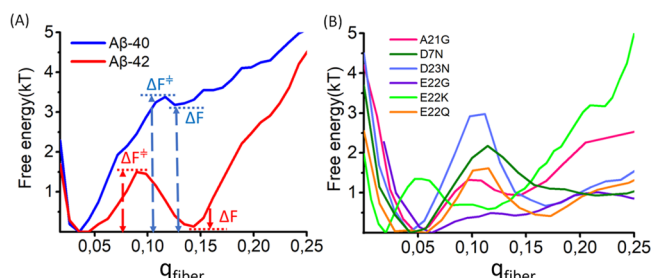


Figure 3. (A) Free energy profile as a function of q_{fiber} for the $A\beta$ -40 and $A\beta$ -42. (B) Free energy profile as a function of q_{fiber} for the $A\beta$ -40 mutants.

energetic difference between extended and more fiberlike structures for the $A\beta$ -40, which is less prominent for the $A\beta$ -42 and the mutants. $A\beta$ -40 presents an almost continuously increasing free energy as a function of increasing q_{fiber} , while the other systems seem to have a more two-state-like behavior. Table 1 shows the values for the free energy profiles, with the

Table 1. Free Energies

species	ΔF (kT)	$\Delta F^\ddagger$ (kT)
$A\beta$ -40	3.2	3.7
$A\beta$ -42	0.2	1.6
A21G	0.9	1.3
D7N	0.9	2.1
D23N	0.9	2.9
E22G	0.8	0.9
E22K	1.1	2.5
E22Q	0.4	1.5

$\Delta F^\ddagger$ indicating the value of the energetic barrier, while the ΔF indicates the height of the first free energy minimum state. Despite having a high free energy cost for fiberlike ordering, the D23N mutation behaves like the $A\beta$ -42, as will be discussed below.

We next analyze the distribution of the structures in the ELViM projection, by reweighing the conformations using the density of microstates $\Omega(E)$ obtained from the WHAM.⁵⁶ Figure 4 shows the result in a gridlike projection with each bin's color indicating the free energy of the ensemble of structures contained within. The region of the α -helix in Figure 1 corresponds to the lowest free energy structural cluster for the $A\beta$ -40, as is pointed out with the red "G" in the upper part of Figure 4A, while the more fiberlike region identified by Figure 4B is highlighted with the black ellipse. The red and yellow lines indicate proposed routes for reconfiguration (the lowest free energy path) when $A\beta$ -40 transverts to assume the fiberlike conformations, with Figure 5 illustrating the most

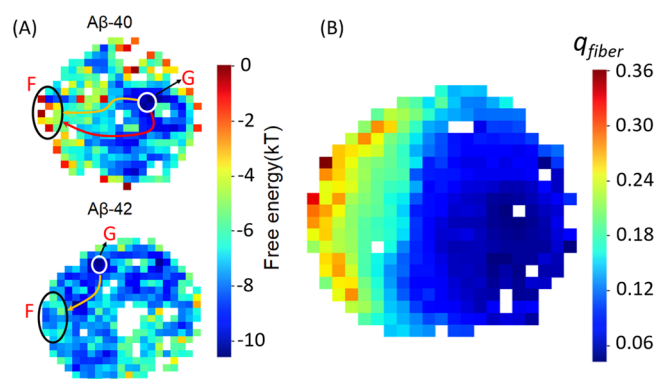
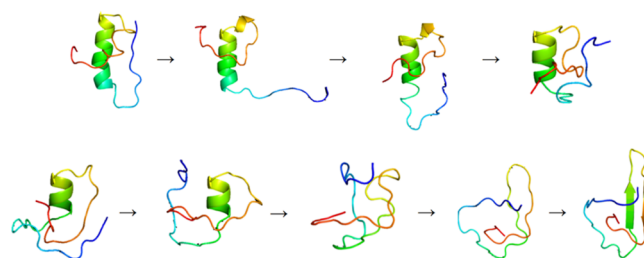


Figure 4. (A) Free energy distribution of the $A\beta$ -40 (upper) and $A\beta$ -42 (lower) structures in the ELViM projections, with the lowest energy region indicated by the "G" and the region of high q_{fiber} indicated by the ellipse with the "F". (B) Mean q_{fiber} distribution of Figure 2.

AB40: Red Path



AB40: Yellow Path

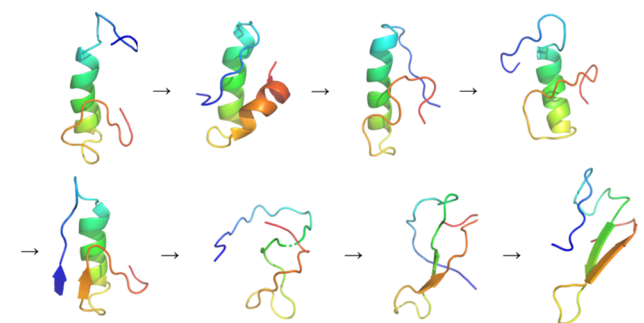


Figure 5. Conformational changes of the structures contained in the possible paths of $A\beta$ -40 to morph into fiberlike structures.

probable conformational changes along each path. The lower part of Figure 4A exhibits the grid projection for the $A\beta$ -42, in which the free energy profile indicates that the lowest free energy structures are nearer the fiber region than for the $A\beta$ -40 case. The most likely pathway of reconfiguration (yellow line) is shorter and more direct, implying that the most stable conformations already have some of the contacts necessary to form a fiber. Figure 6 displays the most representative structures of this path. Video animations showing the conformational changes along the possible routes of Figures 5 and 6 are shown in the SI. The locations of the lowest free energy for both isoforms confirm that the free energy gap between these states and the aggregation-prone states is smaller for $A\beta$ -42 than for $A\beta$ -40, agreeing with Figure 3. This suggests that in some way the aggregation propensity for each peptide may already be encoded in the monomer free energy spectrum.⁵⁷ The same can be seen for the $A\beta$ -40 mutants in

AB42 Path

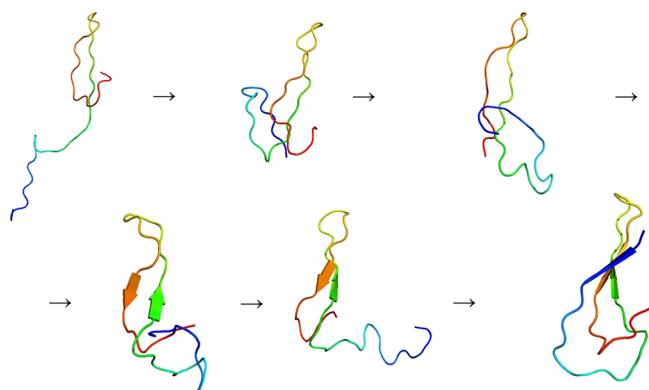


Figure 6. Conformational changes of the structures contained in the possible path of A β -42 to morph into fiberlike structures.

Figure 7, with all low free energy states located in the same region as A β -42 and, as the yellow paths show, following the behavior similar to A β -42.

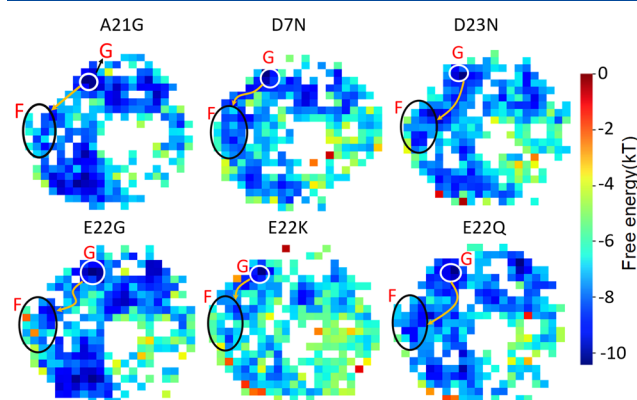


Figure 7. Free energy distribution for each A β -40 mutation with the pathway of lowest free energy highlighted in yellow.

Based on this observation, we calculated the differences in the contact maps averaged over the structures located in the lowest free energy bin and those averaged over structures in the fiber region. For this analysis, we removed the last two residues from A β -42, so we could compare the maps between the species. Figure 8 shows these differences for the A β -40, A β -42, and the mutation E22Q. Apparently, the A β -40 molecule must first undo the α -helix to begin the formation of the first β -sheet contacts, while A β -42 instead undergoes a chain-sliding

mechanism, visible in the parallel contacts between residues 15 and 30, to achieve the fiberlike structure. The E22Q mutation does not have the entire β -sheet formed but also does not exhibit the α -helix contacts and thus does not need to overcome a high energetic barrier to form the fiber, as seen in Figure 3. The contact maps, along with the ELViM projections, confirm that the A β -42 has a higher propensity to form β -strands than A β -40 and its mutants, in harmony with earlier reports.⁵⁸ The differences in the contact maps for the others mutations are available in the Supporting Information in Figure S2, along with the contact maps for each species (Figure S1).

CONCLUSION

The disordered nature and correspondingly large structural fluctuations of IDPs lead to a complex high-dimensional energy landscape. The lack of a funnel to a single dominant structure poses a significant challenge for analysis. Unlike for globular protein folding, it is difficult to determine adequate reaction coordinates. Nevertheless, ELViM allows one to visualize the high-dimensional phase space in an easy-to-understand 2D projection, without the need for a native conformation or a specific reaction coordinate. It is thus a powerful tool for the study of IDPs. In the present work, ELViM permitted the exploration of the conformational changes in A β in its monomeric forms. This study suggests that the aggregation rate is correlated with the β -strand content of the monomers, since the A β -42 and the mutants have higher aggregation rates and exhibit a lower free energy cost to form the β -sheet than the more soluble A β -40. Moreover, the low free energy regions in the projection of the ensembles show that most of the contacts necessary to start the aggregation have already been formed in the A β -42 monomers and the monomers of the less soluble mutants, while the monomers of A β -40 need to undo some helical structure first.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.1c08525>.

Mean contact maps for different regions in the ELViM phase space for each mutant and their relative differences (PDF)

Videos of conformational changes to morph from the lowest free energy region to fiberlike structures through different paths: A β -40 red path (MP4)

A β -40 yellow path (MP4)

A β -42 path (MOV)

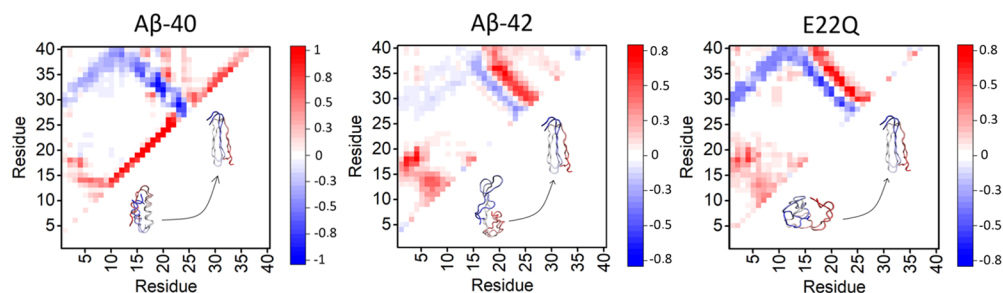


Figure 8. Contact maps. Difference between the mean contact maps for the regions of lowest free energy (red/positive values) and those from the regions of high q_{fiber} (blue/negative values) for the A β -40, A β -42, and E22Q structures.

AUTHOR INFORMATION

Corresponding Author

Vitor B. P. Leite — Department of Physics, Institute of Biosciences, Humanities and Exact Sciences, São Paulo State University (UNESP), São José do Rio Preto, São Paulo 15054-000, Brazil; orcid.org/0000-0003-0008-9079; Email: vitor.leite@unesp.br

Authors

Murilo N. Sanches — Department of Physics, Institute of Biosciences, Humanities and Exact Sciences, São Paulo State University (UNESP), São José do Rio Preto, São Paulo 15054-000, Brazil; orcid.org/0000-0001-9650-7989

Kaitlin Knapp — Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States

Antonio B. Oliveira Jr. — Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States

Peter G. Wolynes — Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States; orcid.org/0000-0001-7975-9287

José N. Onuchic — Center for Theoretical Biological Physics and Departments of Physics and Astronomy, Chemistry, and Biosciences, Rice University, Houston, Texas 77005, United States; orcid.org/0000-0002-9448-0388

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jpcb.1c08525>

Author Contributions

[†]M.N.S. and K.K. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

M.N.S. was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Grant 130147/2020-6). This research was supported by the Center for Theoretical Biological Physics sponsored by the NSF (Grants PHY-2019745). J.N.O. is also supported by the NSF (Grant CHE-1614101) and by the Welch Foundation (Grant C-1792). J.N.O. is a Cancer Prevention and Research Institute of Texas (CPRIT) Scholar in Cancer Research. A.B.O.J. acknowledges the Robert A. Welch Postdoctoral Fellow program. P.G.W. is also supported by the D.R. Bullard-Welch Chair at Rice University (Grant C-0016). V.B.P.L. was supported by CNPq (Grant 310017/2020-3) and FAPESP (Grants 2019/22540-3 and 2018/18668-1).

REFERENCES

- (1) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167–195.
- (2) Wolynes, P.; Luthey-Schulten, Z.; Onuchic, J. Fast-folding eriments and the topography of protein folding energy landscapes. *Chem. Biol.* **1996**, *3*, 425–432.
- (3) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545–600.
- (4) Onuchic, J. N.; Socci, N. D.; Luthey-Schulten, Z.; Wolynes, P. G. Protein folding funnels: the nature of the transition state ensemble. *Folding Des.* **1996**, *1*, 441–450.
- (5) Wolynes, P. G. Folding funnels and energy landscapes of larger proteins within the capillarity approximation. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 6170–6175.
- (6) Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; et al. Intrinsically disordered protein. *J. Mol. Graphics Modell.* **2001**, *19*, 26–59.
- (7) Dyson, H. J.; Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208.
- (8) Kulkarni, P.; Uversky, V. N. Intrinsically Disordered Proteins: The Dark Horse of the Dark Proteome. *Proteomics* **2018**, *18*, No. e1800061.
- (9) Oldfield, C. J.; Dunker, A. K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **2014**, *83*, 553–584.
- (10) Peng, Z.; Yan, J.; Fan, X.; Mizianty, M. J.; Xue, B.; Wang, K.; Hu, G.; Uversky, V. N.; Kurgan, L. Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. *Cell. Mol. Life Sci.* **2015**, *72*, 137–151.
- (11) Jia, Z.; Beugelsdijk, A.; Chen, J.; Schmit, J. D. The Levinthal problem in amyloid aggregation: Sampling of a flat reaction space. *J. Phys. Chem. B* **2017**, *121*, 1576–1586.
- (12) Fisher, C. K.; Stultz, C. M. Constructing ensembles for intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2011**, *21*, 426–431.
- (13) Chebaro, Y.; Ballard, A. J.; Chakraborty, D.; Wales, D. J. Intrinsically disordered energy landscapes. *Sci. Rep.* **2015**, *5*, 1–12.
- (14) Chong, S.-H.; Ham, S. Folding free energy landscape of ordered and intrinsically disordered proteins. *Sci. Rep.* **2019**, *9*, 1–9.
- (15) Jia, Z.; Schmit, J. D.; Chen, J. Amyloid assembly is dominated by misregistered kinetic traps on an unbiased energy landscape. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 10322–10328.
- (16) Uversky, V. N. In *Dancing protein clouds: Intrinsically disordered proteins in health and disease, Part A*; Uversky, V. N., Ed.; Progress in Molecular Biology and Translational Science; Academic Press, 2019; Vol. 166, pp 1–17.
- (17) Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331.
- (18) Van Der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589–6631.
- (19) Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 8183–8188.
- (20) Papoian, G. A. Proteins with weakly funneled energy landscapes challenge the classical structure–function paradigm. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 14237–14238.
- (21) Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. Biomolecular dynamics: order–disorder transitions and energy landscapes. *Rep. Prog. Phys.* **2012**, *75*, 076601.
- (22) Levine, Z. A.; Larini, L.; LaPointe, N. E.; Feinstein, S. C.; Shea, J.-E. Regulation and aggregation of intrinsically disordered peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 2758–2763.
- (23) Duong, V. T.; Chen, Z.; Thapa, M. T.; Luo, R. Computational studies of intrinsically disordered proteins. *J. Phys. Chem. B* **2018**, *122*, 10455–10469.
- (24) Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **2002**, *416*, S07–S11.
- (25) Iadanza, M. G.; Jackson, M. P.; Hewitt, E. W.; Ranson, N. A.; Radford, S. E. A new era for understanding amyloid structures and disease. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 755–773.
- (26) Chiti, F.; Dobson, C. M. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **2017**, *86*, 27–68.

- (27) Jucker, M.; Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **2013**, *501*, 45–51.
- (28) Frieden, C. Protein aggregation processes: In search of the mechanism. *Protein Sci.* **2007**, *16*, 2334–2344.
- (29) Sissoko, C. A. T. *Profiling tau hyperphosphorylation in formalin-fixed paraffin embedded human brain with liquid chromatography tandem mass spectrometry*. PhD thesis, Icahn School of Medicine at Mount Sinai, 2020.
- (30) Haass, C.; Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β -peptide. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 101–112.
- (31) Kaye, R.; Head, E.; Thompson, J. L.; McIntire, T. M.; Milton, S. C.; Cotman, C. W.; Glabe, C. G. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **2003**, *300*, 486–489.
- (32) Copani, A. The underexplored question of β -amyloid monomers. *Eur. J. Pharmacol.* **2017**, *817*, 71–75.
- (33) Ferrone, F. A. Assembly of $A\beta$ proceeds via monomeric nuclei. *J. Mol. Biol.* **2015**, *427*, 287–290.
- (34) Nag, S.; Sarkar, B.; Bandyopadhyay, A.; Sahoo, B.; Sreenivasan, V. K.; Kombrabail, M.; Muralidharan, C.; Maiti, S. Nature of the amyloid- β monomer and the monomer-oligomer equilibrium. *J. Biol. Chem.* **2011**, *286*, 13827–13833.
- (35) Thu, T. T. M.; Co, N. T.; Tu, L. A.; Li, M. S. Aggregation rate of amyloid beta peptide is controlled by beta-content in monomeric state. *J. Chem. Phys.* **2019**, *150*, 225101.
- (36) Zheng, W.; Tsai, M.-Y.; Wolynes, P. G. Comparing the aggregation free energy landscapes of amyloid beta (1–42) and amyloid beta (1–40). *J. Am. Chem. Soc.* **2017**, *139*, 16666–16676.
- (37) Barz, B.; Liao, Q.; Strodel, B. Pathways of amyloid- β aggregation depend on oligomer shape. *J. Am. Chem. Soc.* **2018**, *140*, 319–327.
- (38) Zheng, W.; Tsai, M.-Y.; Chen, M.; Wolynes, P. G. Exploring the aggregation free energy landscape of the amyloid- β protein (1–40). *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 11835–11840.
- (39) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.
- (40) Tomaselli, S.; Esposito, V.; Vangone, P.; van Nuland, N. A.; Bonvin, A. M.; Guerrini, R.; Tancredi, T.; Temussi, P. A.; Picone, D. The α -to- β conformational transition of Alzheimer's $A\beta$ -(1–42) peptide in aqueous media is reversible: a step by step conformational analysis suggests the location of β conformation seeding. *Chem-BioChem* **2006**, *7*, 257–267.
- (41) Vivekanandan, S.; Brender, J. R.; Lee, S. Y.; Ramamoorthy, A. A partially folded structure of amyloid-beta (1–40) in an aqueous environment. *Biochem. Biophys. Res. Commun.* **2011**, *411*, 312–316.
- (42) Attar, A.; Meral, D.; Urbanc, B.; Bitan, G. *Bio-nanoimaging*; Elsevier, 2014; pp 429–442.
- (43) Hatami, A.; Monjazeb, S.; Milton, S.; Glabe, C. G. Familial Alzheimer's disease mutations within the amyloid precursor protein alter the aggregation and conformation of the amyloid- β peptide. *J. Biol. Chem.* **2017**, *292*, 3172–3185.
- (44) Oliveira, A. B., Jr.; Yang, H.; Whitford, P. C.; Leite, V. B. Distinguishing biomolecular pathways and metastable states. *J. Chem. Theory Comput.* **2019**, *15*, 6482–6490.
- (45) Oliveira, A. B., Jr.; Lin, X.; Kulkarni, P.; Onuchic, J. N.; Roy, S.; Leite, V. B. P. Exploring Energy Landscapes of Intrinsically Disordered Proteins: Insights into Functional Mechanisms. *J. Chem. Theory Comput.* **2021**, *17*, 3178–3187.
- (46) Hardin, C.; Eastwood, M. P.; Luthey-Schulten, Z.; Wolynes, P. G. Associative memory Hamiltonians for structure prediction without homology: alpha-helical proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 14235–14240.
- (47) Davtyan, A.; Schafer, N. P.; Zheng, W.; Clementi, C.; Wolynes, P. G.; Papoian, G. A. AWSEM-MD: protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing. *J. Phys. Chem. B* **2012**, *116*, 8494–8503.
- (48) Beebe, R.; Pell, R.; Seasholtz, M. *Chemometrics: A Practical Guide*; Wiley-Interscience: New York, NY, 1998.
- (49) J.N. Miller, J. M. *Statistics and Chemometrics for Analytical Chemistry*; Prentice Hall, 2010.
- (50) Ragonnet-Cronin, M.; Hodcroft, E.; Hué, S.; Fearnhill, E.; Delpech, V.; Leigh Brown, A.; Lycett, S. UK HIV Drug Resistance Database, Automated analysis of phylogenetic clusters. *BMC Bioinformatics* **2013**, *14*, 317.
- (51) Trevor, F.; Cox, M. A. A. C. *Multidimensional Scaling*; Taylor & Francis, 2000.
- (52) France, S.; Carroll, J. Two-Way Multidimensional Scaling: A Review. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* **2011**, *41*, 644–661.
- (53) Tejada, E.; Minghim, R.; Nonato, L. G. On Improved Projection Techniques to Support Visual Exploration of Multi-Dimensional Data Sets. *Information Visualization* **2003**, *2*, 218–231.
- (54) Oliveira, A. B., Jr.; Fatore, F. M.; Paulovich, F. V.; Oliveira, O. N., Jr.; Leite, V. B. P. Visualization of Protein Folding Funnels in Lattice Models. *PLoS One* **2014**, *9*, No. e100861.
- (55) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (56) Sun, L.; Noel, J. K.; Sulkowska, J. I.; Levine, H.; Onuchic, J. N. Connecting thermal and mechanical protein (un) folding landscapes. *Biophys. J.* **2014**, *107*, 2950–2961.
- (57) Chakraborty, D.; Straub, J. E.; Thirumalai, D. Differences in the free energies between the excited states of $A\beta$ 40 and $A\beta$ 42 monomers encode their aggregation propensities. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 19926–19937.
- (58) Côté, S.; Derreumaux, P.; Mousseau, N. Distinct morphologies for amyloid beta protein monomer: $A\beta$ 1–40, $A\beta$ 1–42, and $A\beta$ 1–40 (D23N). *J. Chem. Theory Comput.* **2011**, *7*, 2584–2592.