Assembly of Huntingtin headpiece into α -helical bundles

Beytullah Ozgur

College of Engineering, Koc University, Istanbul 34450, Turkey

Mehmet Sayar^{a)}

Chemical and Biological Engineering and Mechanical Engineering Departments, College of Engineering, Koc University, 34450 Istanbul, Turkey

(Received 22 March 2017; accepted 10 May 2017; published 24 May 2017)

Protein aggregation is a hallmark of neurodegenerative disorders. In this group of brain-related disorders, a disease-specific “host” protein or fragment misfolds and adopts a metastatic, aggregate-prone conformation. Often, this misfolded conformation is structurally and thermodynamically different from its native state. Intermolecular contacts, which arise in this non-native state, promote aggregation. In this regard, understanding the molecular principles and mechanisms that lead to the formation of such a non-native state and further promote the formation of the critical nucleus for fiber growth is essential. In this study, the authors analyze the aggregation propensity of Huntingtin headpiece (*htt*^{NT}), which is known to facilitate the polyQ aggregation, in relation to the helix mediated aggregation mechanism proposed by the Wetzel group. The authors demonstrate that even though *htt*^{NT} displays a degenerate conformational spectrum on its own, interfaces of macroscopic or molecular origin can promote the α -helix conformation, eliminating all other alternatives in the conformational phase space. Our findings indicate that *htt*^{NT} molecules do not have a strong orientational preference for parallel or antiparallel orientation of the helices within the aggregate. However, a parallel packed bundle of helices would support the idea of increased polyglutamine concentration, to pave the way for cross- β structures. © 2017 American Vacuum Society. [<http://dx.doi.org/10.1116/1.4984009>]

I. INTRODUCTION

Expansion of the glutamine tract is the major cause of nine lethal neurological disorders that are collectively called as polyglutamine (polyQ) diseases. Postmortem analysis of brain tissues reveals β -sheet rich structures, a common property of all neurological disorders, in the form of neuronal inclusions.^{1,2} Recent studies indicate that abnormal neuronal activity precedes the cell loss while the site and mechanism of degeneration is specific to the disease itself.³ A characteristic cross- β structure is observed in all polyglutamine fibrils.^{4,5} A major distinction between polyQ peptides and other amyloids is the nature of this cross- β structure. The majority of amyloid peptides favor parallel β -sheets⁶ whereas antiparallel β -sheets are more common in polyQ fibrils.^{7–11}

The presence of an expanded glutamine tract is the only common feature of polyQ peptides. Thus, it was long thought that disease pathology is governed by polyQ tract alone. Studies emphasizing on simple polyQs (individual polyQ tracts) provide invaluable insights regarding the nature of polyQ fibrils, critical length-dependence, and disease onset. The circular dichroism spectra of polyQ peptides with various lengths suggest that simple polyQs exist predominantly as random coils¹² in solution. The aggregation of polyQ peptides is known to follow a nucleated growth mechanism,¹³ and the nucleus size is inversely correlated with glutamine repeat length. The study of Kar *et al.*¹⁴ which systematically examines the relationship between the glutamine repeat length and nucleus size demonstrates that expanding

the glutamine tract from Q_{18} to Q_{37} decreases the nucleus size from 4 to 1 molecule. A monomeric nucleus implies that mutant polyQ peptides, the ones with an excessive number of glutamine repeats ($Q_N > 37$), undergo a coil-to- β -strand transition. There are several structures proposed to explain the underlying transition. Perutz's β -helix structure is one of the earliest models proposed to interpret the available x-ray diffraction data.¹⁵ The stability of β -helix structure strictly depends on the glutamine repeat length and is a rare entity for polyQ peptides with a tract length that is below the critical threshold. Perutz suggested that polyQ fibrils are water-filled nanotubes that involve β -helices as building blocks. The stability and length dependence of the β -helix structure was confirmed by a molecular dynamics study.¹⁶ However, as opposed to the proposed aqueous pore, it was observed that water is strictly excluded from the core. Two other plausible models, β -arc^{9,17–19} and β -turn structures,^{7,14} rely on antiparallel β -hairpins. Recent experiments on polyQ fibrils that aim to distinguish these two models^{20–22} indicate that a β -turn motif fits better to the available data.

With respect to its position in the associated proteins, polyglutamine tract is surrounded by a mixed amino acid sequence that influences the rate and mechanism of polyQ aggregation.²³ Studies on Huntingtin (*htt*) protein best explains this phenomenon. Full-length *htt* is composed of 3144 residues, is localized to cytoplasm^{24,25} and is subjected to heavy enzymatic cleavage.^{26–28} Although the exact structure of *htt* in neuronal aggregates is unknown, polyQ containing *htt* exon I is the smallest physiological N-terminal fragment, which is capable of inducing polyQ aggregation.²⁸ Studies on truncated N-terminal fragments with an expanded

^{a)}Electronic mail: msayar@ku.edu.tr

polQ tract reveal nuclear localization and aggregation.^{29,30} The polyglutamine tract in htt exon 1 is flanked by two sequences; a 17 residue long N-terminus headpiece, known as Huntingtin headpiece (*htt*^{NT}) and a proline rich C-terminus. The x-ray structure of monomeric htt exon-1 determined by fusing to maltose-binding protein³¹ implies an α -helix for *htt*^{NT}. The polyglutamine tract can be unstructured or partly incorporated to the α -helix formed by *htt*^{NT} while the proline-rich C-terminus adopts a polyproline II (PPII) helix. Systematic investigations on simple versus complex polyQ (with the flanking residues) peptides demonstrate an increase in the aggregation rates when polyglutamine tract is flanked by its neighbors.^{32–35} Interestingly, flanking sequences bear opposing roles in polyQ aggregation such that *htt*^{NT} facilitates the aggregation³² whereas the proline-rich C-terminus exhibits a suppressing role³⁶ without altering the aggregation mechanism. The protective role of proline-rich C-terminus is thought to be related to its characteristic PPII-helix structure that reduces the β -sheet propensity of neighboring polyQ tract.^{33,37}

As noted before, *htt*^{NT} which functions as a localization signal,³⁸ is known to accelerate the polyQ aggregation.³² A more intriguing fact is that isolated *htt*^{NT} tracts can form helix-rich globular aggregates.¹⁰ Fusing *htt*^{NT} sequence with an expanded polyglutamine tract ensures that these globular aggregates proceed to high-order, fibrillar structures that are rich in β -sheet content.³⁴ Interestingly, a significant portion of *htt*^{NT} preserves its α -helix content in htt exon 1 fibrils¹⁰ in contrast to the studies³³ that imply a length-dependent incorporation of *htt*^{NT} into β -sheets. Further TEM and AFM analyses of polyQ fibrils concluded that incorporation of a helical domain next to polyQ tract does not disrupt the cross- β -structure.³⁹ Based on experimental evidence, Wetzel Group proposed an *htt*^{NT} mediated polyQ aggregation mechanism²³ (also known as the proximity model). This mechanism suggests a reversible association of htt exon 1 fragments into tetrameric and oligomeric assemblies induced by the self-interactions of *htt*^{NT} tract. Upon bundle formation *htt*^{NT} molecule adopts an α -helix and these helix-rich bundles lead to an increase in the local concentration of polyglutamine tracts. Interactions between neighboring polyglutamine tracts induce a coil-to- β -sheet transition in this tract that is capable of nucleating further fibril growth.

Wetzel's proposed mechanism suggests that early events of polyQ aggregation is guided by *htt*^{NT} tract alone. The role of helix-rich oligomers as on-pathway intermediates have been reported for several systems including α -synuclein,⁴⁰ Islet amyloid polypeptide,⁴¹ and A β .⁴² However, *htt*^{NT} is unique in the sense that in other systems helix-rich assemblies disappear as aggregation proceeds whereas *htt*^{NT} preserves its helical character in final aggregates. In addition, the fact that isolated *htt*^{NT} tracts can form globular aggregates draws a lot of interest. *htt*^{NT} sequence possesses an amphiphilic character which is found to be exploited by chaperonin TRiC, a known inhibitor of htt aggregation, that sequesters *htt*^{NT} by recognizing its hydrophobic residues. Mishra *et al.* also makes use of this amphiphilic character to develop novel inhibitory

molecules.⁴³ Two mutations on *htt*^{NT} tract, S13D and S16D, are also known to have an inhibitory effect on htt aggregation⁴⁴ and are thought to cause a loss of cross-talk between *htt*^{NT} and the polyQ tract.⁴⁵ *htt*^{NT} is also known to regulate the aggregation of htt exon 1 fragments on lipid membranes.^{46–48} To explain the mechanism of htt exon 1 aggregation on lipid membranes Michalek *et al.*⁴⁹ proposed a model analogous to the Wetzel's proximity model where the amphiphatic nature of *htt*^{NT} guides the membrane anchorage and aggregation.

Experimental studies on isolated *htt*^{NT} molecules demonstrate that *htt*^{NT} monomers in solution are disordered.^{32,34} On the other hand, an atomistic characterization of *htt*^{NT} through computational methods suggests that *htt*^{NT} mostly populates a two helix bundle with hydrophobic residues facing the bundle core.⁵⁰ The combined study of Williamson *et al.*⁵¹ reports an inverse correlation between the helix content of *htt*^{NT} and polyglutamine tract length. Interestingly, the authors report that *htt*^{NT} suppresses the aggregation propensity of polyQ tract. A recent characterization⁵² of the free energy surface of *htt*^{NT} on its own and when fused to accompanying htt exon 1 sequences suggests that isolated *htt*^{NT} tract is mostly disordered with a decent α -helix propensity in N-terminus. Rarely, salt-bridges between oppositely charged side chains are formed. Fusing polyQ tract (*Q*₁₇) shifts the α -helix propensity of *htt*^{NT} toward the C-terminus which can sequester the neighboring glutamine residues as well. Also, interactions between the glutamine tract and nonpolar side chains of *htt*^{NT} lead to a decrease in the solvent accessible surface area (SASA) of the latter. A further addition of proline-rich C-terminus greatly enhances the α -helix propensities of both *htt*^{NT} and polyglutamine tracts and nonpolar side chains of *htt*^{NT} tract becomes solvent exposed while polar ones interact with the glutamine tract.⁵² A solvent exposed *htt*^{NT} for nonpathogenic htt exon 1 fragments contradicts with the computational findings of Dlugosz and Trylska⁵³ which suggests nonpolar contacts between *htt*^{NT} and proline-rich C-terminus. Such nonpolar contacts are protective as they prevent *htt*^{NT} from becoming solvent exposed. However, expanding the glutamine tract leads to a loss of these nonpolar contacts with *htt*^{NT} becoming solvent exposed and prone to self-association.⁵³

In the context of Wetzel Group's proximity model, the self-association of *htt*^{NT} is an essential step. Although the aggregation propensity of *htt*^{NT} is implied by several studies, the identity of relevant forces that guide its assembly into multimeric assemblies is not understood well. To the best of our knowledge, the study of Trylska and Dlugosz⁵³ which performs a limited analysis on *htt*^{NT} dimers is the only study that examines the aggregation propensity of *htt*^{NT} in an aggregate level. Here, through microsecond-long molecular dynamics simulations, we systematically examine the conformational spectrum of *htt*^{NT} (Fig. 1) in macroscopic and microscopic interfaces with a particular focus on the interplay between three main forces; Coulomb interactions, hydrophobic forces and hydrogen bonding which are encoded in the *htt*^{NT} sequence.

FIG. 1. Helical wheel representation of the 17 residue htt^{NT} tract with amino acid sequence *MATLEKLMKAFESLKSF*. htt^{NT} displays a strong secondary amphiphilic character with hydrophobic/hydrophilic residues strictly partitioned on opposite sides of the helix. Hydrophobic and hydrophilic residues are shown with diamonds and circles, whereas positively and negatively charged residues are shown with pentagons and triangles, respectively. Shades of green and orange denote the degree of hydrophobicity/hydrophilicity (stronger the character darker the color).

II. METHODS

Molecular dynamics simulations on htt^{NT} peptide were carried out in GROMACS 4.5.6 (Ref. 54) by using GROMOS54a7 force field.⁵⁵ Water was represented by SPC-E water model.⁵⁶ Twin-range cutoff scheme was used to calculate the non-bonded interactions with a cutoff distance of 1 nm for neighbor list (*rlist*) and 1.4 nm for VdW interactions (*rvdw*). The neighbor list was updated every ten steps and long range dispersion corrections were applied for energy. Coulomb interactions were calculated by particle mesh Ewalds (PME) method⁵⁷ with a real space cutoff distance of 1 nm (*rcoulomb*) and a Fourier grid spacing of 0.12 nm. Linear constraint solver algorithm⁵⁸ was used to restrain all-bonds, which allowed a time step of 2 fs. Simulations were conducted in the canonical ensemble, where the temperature was kept constant at 298 K with the velocity rescaling algorithm⁵⁹ with a time constant of 0.1 ps (*tau_t*).

The N- and C-terminus of htt^{NT} molecule were capped with acetyl (ACE) and NH_2 , respectively. htt^{NT} contains three lysine and two glutamic acid residues. Two Na and three Cl ions were added per molecule to the simulation box to have a charge neutral system. Prior to the MD simulations of htt^{NT} molecules, all peptide atoms were fixed for 600 ps allowing solvent to relax. The interface simulations of htt^{NT} dimers involved two extended htt^{NT} molecules placed close to each other. The following routine was used to set up this initial structure. First, we fixed the end-to-end distance of an htt^{NT} molecule and performed a 20 ns-long simulation at the air/water interface. This allowed a proper partitioning of hydrophobic/hydrophilic

groups without disrupting the secondary structure of the peptide. Finally, the structure obtained at the end of 20 ns was duplicated with both molecules lying at the air/water interface. This initial structure was equilibrated through a series of 600 ps simulations where restraints on htt^{NT} molecules were gradually removed.

The SASA of hydrophobic side chains of htt^{NT} molecule(s) was calculated by the *g_sas* tool of GROMACS with a probe size of 0.14 nm. The choice of hydrophobic residues for SASA calculations rely on the Kyte-Doolittle⁶⁰ hydrophobicity scale, according to which methionine, leucine, phenylalanine, and alanine are ranked as hydrophobic. For the bulk water simulations, the surface and output groups for *g_sas* tool were chosen as the peptide and side chains of the hydrophobic residues. For SASA calculations at the air/water interface, we first calculated the SASA of hydrophobic groups by choosing the peptide and hydrophobic side chains as surface and output groups, respectively. Let us say S_1 denotes the SASA obtained by this calculation. Next, we changed the choice of surface group from peptide to the whole system, including water, which yields the SASA of hydrophobic side chains that actually face the air. Let us say S_2 denotes the SASA obtained by the second calculation. The actual SASA of the hydrophobic side chains that were exposed to water was found simply by calculating the difference: $S_1 - S_2$.

In order to compare intra- and intermolecular Coulomb interactions between charged side chains, water and counterions were removed from the simulation box. The dielectric constant was set to 80, and PME was replaced with the simple cutoff method (with the cutoff distance set as half the minimum box length). Using this set of parameters, the Coulomb energies of all simulations were recalculated using the *rerun* option of *mdrun* tool in GROMACS. This calculation is used merely to compare the distributions of the charged residues within the aggregates.

In order to determine the orientational flexibility of htt^{NT} dimers, the N-terminus fragment spanning the region from Thr3 to Ala10, which is observed to exhibit the highest helix propensity, was used to assign a vector to each of the two htt^{NT} molecules. The angle between these two vectors which is equal to 0° for parallel and 180° for antiparallel configurations is calculated to determine the orientational flexibility of htt^{NT} dimers.

For secondary structure assignment define secondary structure of proteins algorithm⁶¹ was used. Molecular graphics were produced by the visual molecular dynamics program,⁶² and plots were produced with the Gnuplot 5.0 package. Even though the simulations were performed with explicit water, in all molecular graphics presented here, water molecules are not shown for visual clarity. Only in Fig. 4, in order to depict the air/water interface, water molecules are shown in blue. A comprehensive list of all simulations discussed in this manuscript are provided in the supplementary material, Sec. I.⁶⁹

III. RESULTS AND DISCUSSION

As shown in Fig. 1, htt^{NT} displays a secondary amphiphilic⁶³ character in the α -helix conformation. In dilute

solutions, where the molecule can be considered to be isolated in bulk water, such a strong partitioning of the hydrophobic/hydrophilic residues leads to the destabilization of this conformation. In order to demonstrate this we set up two different simulations of *htt*^{NT} in bulk water. Microsecond long simulations starting from an α -helix and random coil structure both suggest that residues 3–10 have a strong helical propensity, whereas the remaining residues fail to adopt a stable conformation [Figs. 2(a) and 2(b)]. The SASA analysis for the hydrophobic residues (Fig. 3) demonstrates that both simulations yield an average hydrophobic contact area of $\approx 7.2 \text{ nm}^2$. The molecule attempts to reduce this unfavorable contact area by partial unfolding of the termini (in particular, the C-terminus) of the peptide to cover up the hydrophobic surface of the helical portion as shown with the snapshots of the final structure for both simulations in Fig. 2. The average number of backbone hydrogen bonds (also

provided in Fig. 3) also indicates that roughly half of the molecule remains in helical conformation in dilute solution.

The conformational behavior of *htt*^{NT} closely resembles the LK α 14 peptide with the amino acid sequence Ace-(LKKLLKL)₂-Nme.^{64,65} Similar to *htt*^{NT}, LK α 14 also possesses a large hydrophobic moment⁶³ in the α -helix conformation and lacks a well-defined secondary structure in dilute solution.

The conformational dynamics of LK α 14 is driven by the competition between hydrophobic forces due to the leucine side chains, Coulomb interactions due to the lysine residues, and backbone hydrogen bonding. Unlike LK α 14's large bare charge (+6), *htt*^{NT} has 3 lysine and 2 glutamic acid residues, with an overall charge of +1. Hence, for *htt*^{NT}, the charged side chains do not play a crucial role in formation of the secondary structure in bulk water. The changes in short range Coulomb energy during the 1 μs simulations (see

FIG. 2. Secondary structure of a single *htt*^{NT} in bulk water: simulations with helical (a) and random (b) initial structures display a strong α -helix tendency for residues 3–10, while the other residues mostly remain in a disordered state. Snapshots of initial (left) and final (right) conformations, as well as the DSSP analysis of the 1 μs long simulations are shown for both.

FIG. 3. SASA for hydrophobic side chains and average number of backbone hydrogen bonds for single, dimer, tetramer, and dodecamer simulations in bulk water and at the air/water interface. See supplementary material, Sec. II (Ref. 69) for the timeline data for SASA, number of hydrogen bonds, and in addition the short range Coulomb energy of charged residues.

supplementary material, Sec. II)⁶⁹ are rather small compared to the energetic contributions from hydrophobic forces and backbone hydrogen bonding. When the molecule is isolated in bulk water, the competition between backbone hydrogen bond formation (which favors helix formation) and the hydrophobic forces (which drive the molecule into a compact state) compete. The partial α -helical character observed in our simulations is in line with the previous experimental and computational studies.^{10,34,50–52} One should keep in mind that with microsecond conformational transition times, standard MD simulations can only provide a limited perspective on the full diversity of the conformational phase space of such molecules, and for a better understanding, one should resort to enhanced sampling techniques.^{50,52,53,65}

Despite the conformational heterogeneity in bulk water, in the presence of a macroscopic hydrophobic/hydrophilic interface, molecules with a secondary amphiphilic character display a strong partitioning of the side chains⁶⁶ and thereby eliminate all other competing conformations, adopting a unique secondary structure.^{64,65,67} Not surprisingly, *htt*^{NT} also displays a similar behavior as observed in our simulations at the air/water interface (Fig. 4). In a similar manner to the bulk water simulations, we performed two separate simulations of a single *htt*^{NT} in a slab of water with α -helix and random coil as initial conformations. Snapshots for the initial and final conformations of the peptide at the interface and the DSSP analysis of secondary structure are shown for both simulations in Fig. 4. The simulation with the α -helical initial conformation remains stable throughout the simulation [Fig. 4(a)]. In the simulation with a random initial conformation, the N-terminus of the peptide rapidly adopts the α -helix conformation in line with the helical tendency of this part in bulk water simulations. The remaining segment also attains the α -helix conformation around 350 ns.

Thanks to the secondary amphiphilic character of the molecule, the competition between hydrophobic collapse and backbone hydrogen bonding observed in bulk water is

completely annihilated at the interface. As seen in the timeline data for SASA (supplementary material, Sec. II)⁶⁹ even in the initial random state, the molecule already shields all its hydrophobic residues from water. In the absence of any other opposing forces, the backbone hydrogen bonding drives the system toward a full α -helix. In addition, the short range Coulomb energy also benefits from this transition (supplementary material, Sec. II).⁶⁹

Similar to a macroscopic interface, molecular interfaces could also help stabilize the conformation of such secondary amphiphilic peptides.^{64,65,68} In the case of LK α 14 presence of a second peptide suffices for the formation of a stable dimer with both peptides in the full α -helix conformation. In order to understand the behavior of *htt*^{NT}, we set up a simulation with two peptides in random conformations and initially separated from each other with several layers of water in between [in Fig. 5(a) upper-left corner only one of the peptides is shown]. Similar to the bulk water simulations with a single peptide, both molecules adopt the α -helix conformation on their N-termini [Fig. 5(b) DSSP data] while they are still separated. The center-of-mass distance between the peptides [Fig. 5(c)] shows that the peptides aggregate around 200 ns. This initial aggregation is driven by a hydrophobic collapse as evidenced by the sharp drop in SASA [Fig. 5(c)]. After the aggregation, both peptides increase their α -helix content; however, neither of them transforms into a full helix. One peptide forms a broken α -helix and the other one remains in half- α -helix state. Interestingly, the increase in the α -helix content takes place at the cost of an increase in SASA for the hydrophobic side chains.

In comparison to LK, which attains the full α -helix conformation upon dimer formation, *htt*^{NT} only slightly increases its α -helix content in a dimer. Once again their C-termini unfold, whereas the N-termini remain α -helical. The helical segments of the two peptides form hydrophobic contacts and the hydrophobic side chains on the unfolded portion of the molecules also attach to this hydrophobic core from the sides [see Fig. 5(a) snapshot at 210 ns]. The molecules lower their SASA values to a moderate degree compared to a single molecule in bulk (Fig. 3).

The dimer is rather dynamic as seen in the helix angle versus distance correlation plot [Fig. 5(d)]. Since the peptides are not fully helical, only the helical segment of each molecule is taken into account for determining the helix axis and the center of mass distance. As seen in Fig. 5(d), the two molecules' helical portions remain at an angle of 120°. Unlike the LK α 14 molecule, the dimer formation cannot promote full α -helix conformation nor it creates a stable dimeric structure.

Since standard MD simulations are prone to be kinetically trapped in a local energy minimum, we also performed two additional simulations where a pair of *htt*^{NT} molecules is initially packed with their hydrophobic faces in contact in parallel and antiparallel arrangements. DSSP analysis and the distance versus angle correlation plots for these simulations are shown in Fig. 6. The average SASA of both simulations are lower than the random dimer simulation case (Fig. 3),

FIG. 4. Secondary structure of a single *htt*^{NT} at the air/water interface. The molecule adopts the α -helix conformation independent of the initial structure [α -helix in (a) and random coil (b)]. Strong partitioning of the hydrophobic/hydrophilic residues enforces the α -helix conformation for all but the terminal residues as shown in the final conformation of the molecules in both simulations (snapshots on the right).

and the backbone hydrogen bond numbers are also slightly higher, suggesting that the simulation with the random initial conformation might be kinetically trapped in a local minimum. When compared to the single molecule at the air/water interface, though, helix stabilizing effect of these molecular interfaces is not as strong as the macroscopic interface. Between the two arrangements, the parallel one displays a slightly higher stability as evidenced by the bright spot in the correlation plot in Fig. 6(c), as well as the lower SASA and higher backbone hydrogen bond numbers in Fig. 3.

In the absence of any opposing forces further growth of *htt*^{NT} aggregates is also feasible. As the system size limits an analysis starting from a randomly dispersed solution beyond two molecules, for larger size aggregates, we started with a preformed aggregate and tested the structural stability of these aggregates. If one considers the tetrameric state, two possible packing arrangements stand out: all parallel and the checker board structure (where neighboring

molecules are in antiparallel arrangement). In Fig. 7, final structures of the tetrameric aggregates at the end of microsecond long simulations are shown. DSSP analysis of the checkerboard aggregate is also shown in Fig. 7(c). Unlike the dimer case, within a tetramer all *htt*^{NT} molecules remain fully stable in the α -helix conformation. The parallel case also displays a similar stability, where the DSSP analysis is provided in supplementary material, Sec. III.⁶⁹ In both cases, on the average, more than 10 backbone hydrogen bonds are formed and per monomer SASA values are very close to the interface case (Fig. 3).

The angle versus distance correlation plots for all four neighboring helices are provided in supplementary material, Sec. III (Ref. 69) for both the parallel and antiparallel cases. For the parallel case, all helices maintain an average distance of 1.10 nm and an angle of approximately 30°. On the other hand, the antiparallel packed aggregate displays two peaks, which suggests a more dynamic structure compared to the

FIG. 5. Assembly of two *htt*^{NT} molecules in bulk water. (a) Initially the peptides are separated and in random conformations (only a single peptide is shown in the upper-left snapshot). Around 0.2 μ s they approach each other and aggregate through their hydrophobic faces. α -helix content of the molecules further increases, however transition to full α -helix does not take place. DSSP secondary structure assignments (b), center of mass distance between the molecules and SASA (c), and center of mass distance vs angle between helix axes (d) are provided.

FIG. 6. Comparison of parallel and antiparallel arrangement for a pair of *htt*^{NT} peptides in bulk water. *htt*^{NT} molecules are in contact via their hydrophobic faces and are initially oriented parallel (a), (c) or antiparallel (b), (d) with respect to each other. For both orientations, the DSSP plot traces the evolution of secondary structures and angle-distance correlation plot shows the orientational flexibility of *htt*^{NT} dimers.

parallel. The phenylalanine residues located at the 11th and 17th positions of the molecule play an important role in this difference. The DSSP analysis (shown in supplementary material, Sec. III)⁶⁹ suggests that for both cases the 17th residue is permanently unfolded. If one compares the radius of gyration for the parallel versus antiparallel cases (supplementary material, Sec. III),⁶⁹ the parallel case allows the unfolded 17th phenylalanine side chains to interact with the 11th phenylalanine residues to enjoy aromatic interactions within the aggregate. For the antiparallel packing arrangement such a local interaction hub for aromatic rings is not

present. However, this increased concentration of aromatic interactions does not translate into a more stable aggregate: For both the parallel and antiparallel cases the radius of gyration of the whole aggregate is roughly the same (supplementary material, Sec. III).⁶⁹

Compared to LK α 14, which is designed to have an ideal amphiphilic character, *htt*^{NT} still has a few residues breaking the perfect partitioning of the hydrophobic/hydrophilic groups in the α -helix conformation. As a result, even though LK α 14 was able to attain full stability of α -helix conformation with only two molecules, for *htt*^{NT} it takes four molecules to

Fig. 7. Final structures of antiparallel (a) and parallel (b) tetramers at the end of microsecond long simulations. In both structures molecules preserve their α -helix conformation as seen in the DSSP analysis (c) (shown only for the antiparallel). Parallel packing allows strong interaction of phenylalanine side chains further stabilizing the structure.

achieve a similar stability. However, unlike LK α 14 (for which growth beyond a tetramer is prevented by the electrostatic repulsion due to lysines), *htt*^{NT} can still grow. As a final demonstration of the enhanced stability upon aggregation, we tested an aggregate of 12 *htt*^{NT} molecules (Fig. 8). Once again, the larger aggregate size both promotes further α -helix stability and also reduces the hydrophobic side chain SASA to bring it closer to the interface shielding (Fig. 3). DSSP analysis of the dodecamer is provided in supplementary material, Sec. IV.⁶⁹

Aggregation propensity of peptides could differ between bulk and interface as demonstrated by LK α 14.⁶⁵ When the lysine side chains fully align at the interface, the repulsion between molecules becomes even stronger, such that even a dimer cannot maintain full contacts. The lack of such a repulsive interaction for *htt*^{NT} suggests that aggregation should be possible also at the interface. In our simulation with two extended molecules at the air/water interface, which were initially aligned parallel, molecules first form interpeptide backbone hydrogen bonds (Fig. 9). However, both molecules transform to an α -helix conformation during a 1.5 μ s simulation as evidenced by the DSSP and backbone hydrogen bonds shown in Fig. 9. The initial parallel orientation of the peptides prolongs to the final stage as well, despite going through several orientations during the folding process.

In order to determine if the molecules can form even larger structures and whether if they have a preferred orientation, we also performed two different simulations with four peptides at the interface. The peptides were initially in an α -helix conformation and aligned at the interface with their hydrophobic faces facing away from water. In the first simulation, all peptides were aligned in parallel, whereas for the second simulation, they were aligned in antiparallel orientation. DSSP analysis and sample snapshots from the final structures are provided in supplementary material, Sec. V.⁶⁹ Both parallel and antiparallel configurations remained stable through the microsecond simulations. Coulomb energies indicate that the parallel arrangement is slightly unfavorable compared to antiparallel. As seen in Fig. 3 they lower the SASA and increase the hydrogen bonds compared to the single molecule at the interface. Hence, we can conclude that identical to its bulk behavior, *htt*^{NT} also aggregates at the interface and can grow into arrays of aligned helices. Even though all four molecules remain in contact during the

Fig. 8. Dodecamer formed by parallel packed *htt*^{NT} molecules also remains stable [see supplementary material, Sec. IV for the DSSP analysis (Ref. 69)].

FIG. 9. Assembly of two *htt*^{NT} molecules at the air/water interface. Two extended *htt*^{NT} molecules, that are initially separated, rapidly align their backbones, convert to α -helices and align parallel (a). DSSP analysis for both molecules (b) and timeline of intra- and interchain hydrogen bonds (c) demonstrate the formation of the α -helices. Water molecules are not shown for visual clarity.

simulation, visual inspection of the trajectories suggest that the molecules form pairs with multiple hydrophobic contacts in between.

IV. CONCLUSION

In this study, by using microsecond long MD simulations, we have analyzed the conformation and aggregation behavior of *htt*^{NT} peptide. This peptide, with a large hydrophobic moment, displays a secondary amphiphilic character in the α -helix conformation. Here, we demonstrate that, as a direct consequence of its amphiphilic nature, a stable α -helix conformation can be enforced on *htt*^{NT} peptide by exposing it to a macroscopic or molecular interface.

In solution, *htt*^{NT} molecules are driven toward each other via hydrophobic forces. This initial hydrophobic collapse is

followed by a rapid conformational transformation to the α -helix conformation in the presence of a macroscopic interface or other peptides which present a hydrophobic/hydrophilic interface. The lack of any opposing forces against aggregation enables *htt*^{NT} to aggregate both in bulk and at the air/water interface. Our analysis indicates that the molecule does not have a strong preference for the parallel or antiparallel arrangement of the neighboring molecules in the aggregate, neither in bulk nor at the air/water interface.

Our findings support the idea of helix mediated aggregation mechanism of polyQ segments as proposed by the Wetzel group.²³ According to this mechanism, α -helix templated *htt*^{NT} aggregation precedes the polyglutamine aggregation. Only within these helix bundles, polyQ segments reach the required local density to transform into the cross beta structure. Since we studied only the *htt*^{NT} block, we cannot directly comment on the behavior of polyQ section of exon 1. However, it is clear that the parallel packing of the molecules inside the aggregate would provide a higher density of polyQ's in comparison to antiparallel arrangement.

The aggregation tendency of the molecule at the interface is also intriguing with regards to membrane activity of the peptide.⁴⁶ Our simulations, clearly show that *htt*^{NT} peptides readily assemble into one-dimensional aggregates at the air/water interface. Such an arrangement, could strongly enhance the surface aggregation propensity of polyQ segments similar to the bulk case.

Comparison of *htt*^{NT} with the synthetic peptide LK α 14 also provides further insights into the driving forces for aggregation. Unlike LK α 14, *htt*^{NT} lacks a strong opposing force for aggregation, and as a result the size of *htt*^{NT} aggregates are not limited by any thermodynamic forces.

ACKNOWLEDGMENT

M. Sayar is thankful to TÜBİTAK (Project No. 116Z512) for the financial support.

¹C. A. Ross, *Neuron* **19**, 1147 (1997).

²M. Arrasate, S. Mitra, E. S. Schweitzer, M. R. Segal, and S. Finkbeiner, *Nature* **431**, 805 (2004).

³R. Chopra and V. G. Shakkottai, *Neurotherapeutics* **11**, 751 (2014).

⁴M. F. Perutz, T. Johnson, M. Suzuki, and J. T. Finch, *PNAS* **91**, 5355 (1994).

⁵E. Scherzinger *et al.*, *Cell* **90**, 549 (1997).

⁶T. L. Benzinger, D. M. Gregory, T. S. Burkoth, H. Miller-Auer, D. G. Lynn, R. E. Botto, and S. C. Meredith, *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13407 (1998).

⁷A. K. Thakur and R. Wetzel, *PNAS* **99**, 17014 (2002).

⁸D. Sharma, L. M. Shinchuk, H. Inouye, R. Wetzel, and D. A. Kirschner, *Proteins* **61**, 398 (2005).

⁹R. Schneider *et al.*, *J. Mol. Biol.* **412**, 121 (2011).

¹⁰V. N. Sivanandam, M. Jayaraman, C. L. Hoop, R. Kodali, R. Wetzel, and P. C. A. van der Wel, *J. Am. Chem. Soc.* **133**, 4558 (2011).

¹¹D. Punihaole, R. J. Workman, Z. Hong, J. D. Madura, and S. A. Asher, *J. Phys. Chem. B* **120**, 3012 (2016).

¹²S. Chen, V. Berthelier, W. Yang, and R. Wetzel, *J. Mol. Biol.* **311**, 173 (2001).

¹³S. Chen, F. A. Ferrone, and R. Wetzel, *PNAS* **99**, 11884 (2002).

¹⁴K. Kar, M. Jayaraman, B. Sahoo, R. Kodali, and R. Wetzel, *Nat. Struct. Mol. Biol.* **18**, 328 (2011).

- ¹⁵M. F. Perutz, J. T. Finch, J. Berriman, and A. Lesk, *PNAS* **99**, 5591 (2002).
- ¹⁶S. D. Khare, F. Ding, K. N. Gwanmesia, and N. V. Dokholyan, *PLoS Comput. Biol.* **1**, e30 (2005).
- ¹⁷P. Sikorski and E. Atkins, *Biomacromolecules* **6**, 425 (2005).
- ¹⁸A. V. Kajava, U. Baxa, R. B. Wickner, and A. C. Steven, *PNAS* **101**, 7885 (2004).
- ¹⁹A. V. Kajava, U. Baxa, and A. C. Steven, *Faseb J.* **24**, 1311 (2010).
- ²⁰K. Kar, C. L. Hoop, K. W. Drombosky, M. A. Baker, R. Kodali, I. Arduini, P. C. A. van der Wel, W. Seth Horne, and R. Wetzel, *J. Mol. Biol.* **425**, 1183 (2013).
- ²¹L. E. Buchanan, J. K. Carr, A. M. Fluitt, A. J. Hoganson, S. D. Moran, J. J. de Pablo, J. L. Skinner, and M. T. Zanni, *PNAS* **111**, 5796 (2014).
- ²²C. L. Hoop *et al.*, *PNAS* **113**, 1546 (2016).
- ²³R. Wetzel, *J. Mol. Biol.* **421**, 466 (2012).
- ²⁴K. E. De Rooij, J. C. Dorsman, M. A. Smoor, J. T. Den Dunnen, and G. J. Van Ommen, *Hum. Mol. Genet.* **5**, 1093 (1996).
- ²⁵J. K. Cooper *et al.*, *Hum. Mol. Genet.* **7**, 783 (1998).
- ²⁶T. Ratovitski *et al.*, *J. Biol. Chem.* **284**, 10855 (2009).
- ²⁷A. T. N. Tebbenkamp, K. W. Crosby, Z. B. Siemienski, H. H. Brown, T. E. Golde, D. R. Borchelt, and X. Jiang Li, *PLoS One* **7**, e50750 (2012).
- ²⁸C. Landles *et al.*, *J. Biol. Chem.* **285**, 8808 (2010).
- ²⁹V. C. Wheeler *et al.*, *Hum. Mol. Genet.* **9**, 503 (2000).
- ³⁰J. Cornett, F. Cao, C.-E. Wang, C. A. Ross, G. P. Bates, S.-H. Li, and X.-J. Li, *Nat. Genet.* **37**, 198 (2005).
- ³¹M. Whi Kim, Y. Chelliah, S. Woo Kim, Z. Otwinowski, and I. Bezprozvanny, *Structure* **17**, 1205 (2009).
- ³²A. K. Thakur *et al.*, *Nat. Struct. Mol. Biol.* **16**, 380 (2009).
- ³³V. V. Lakhani, F. Ding, and N. V. Dokholyan, *PLoS Comput. Biol.* **6**, e1000772 (2010).
- ³⁴M. Jayaraman, R. Kodali, B. Sahoo, A. K. Thakur, A. Mayasundari, R. Mishra, C. B. Peterson, and R. Wetzel, *J. Mol. Biol.* **415**, 881 (2012).
- ³⁵S. L. Crick, K. M. Ruff, K. Garai, C. Frieden, and R. V. Pappu, *PNAS* **110**, 20075 (2013).
- ³⁶A. Bhattacharyya, A. K. Thakur, V. M. Chellgren, G. Thiagarajan, A. D. Williams, B. W. Chellgren, T. P. Creamer, and R. Wetzel, *J. Mol. Biol.* **355**, 524 (2006).
- ³⁷C. L. Hoop, H.-K. Lin, K. Kar, Z. Hou, M. A. Poirier, R. Wetzel, and P. C. A. van der Wel, *Biochemistry* **53**, 6653 (2014).
- ³⁸T. Maiuri, T. Woloshansky, J. Xia, and R. Truant, *Hum. Mol. Genet.* **22**, 1383 (2013).
- ³⁹B. Kokona, Z. P. Rosenthal, and R. Fairman, *Biochemistry* **53**, 6738 (2014).
- ⁴⁰V. L. Anderson, T. F. Ramlall, C. C. Rospigliosi, W. W. Webb, and D. Eliezer, *PNAS* **107**, 18850 (2010).
- ⁴¹G. Liu, A. Prabhakar, D. Aucoin, M. Simon, S. Sparks, K. J. Robbins, A. Sheen, S. A. Petty, and N. D. Lazo, *J. Am. Chem. Soc.* **132**, 18223 (2010).
- ⁴²M. D. Kirkitadze, M. M. Condrón, and D. B. Teplow, *J. Mol. Biol.* **312**, 1103 (2001).
- ⁴³R. Mishra, M. Jayaraman, B. P. Roland, E. Landrum, T. Fullam, R. Kodali, A. K. Thakur, I. Arduini, and R. Wetzel, *J. Mol. Biol.* **415**, 900 (2012).
- ⁴⁴X. Gu *et al.*, *Neuron* **64**, 828 (2009).
- ⁴⁵B. Kokona, K. A. Johnson, and R. Fairman, *Biochemistry* **53**, 6747 (2014).
- ⁴⁶M. Michalek, E. S. Salnikov, and B. Bechinger, *Biophys. J.* **105**, 699 (2013).
- ⁴⁷K. A. Burke, K. J. Kauffman, C. Samuel Umbaugh, S. L. Frey, and J. Legleiter, *J. Biol. Chem.* **288**, 14993 (2013).
- ⁴⁸A. Nagarajan, S. Jawahery, and S. Matysiak, *J. Phys. Chem. B* **118**, 6368 (2014).
- ⁴⁹M. Michalek, E. S. Salnikov, S. Werten, and B. Bechinger, *Biochemistry* **52**, 847 (2013).
- ⁵⁰N. W. Kelley, X. Huang, S. Tam, C. Spiess, J. Frydman, and V. S. Pande, *J. Mol. Biol.* **388**, 919 (2009).
- ⁵¹T. E. Williamson, A. Vitalis, S. L. Crick, and R. V. Pappu, *J. Mol. Biol.* **396**, 1295 (2010).
- ⁵²V. Binette, S. Côté, and N. Mousseau, *Biophys. J.* **110**, 1075 (2016).
- ⁵³M. Długosz and J. Trylska, *J. Phys. Chem. B* **115**, 11597 (2011).
- ⁵⁴S. Pronk *et al.*, *Bioinformatics* **29**, 845 (2013).
- ⁵⁵N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. van Gunsteren, *Eur. Biophys. J.* **40**, 843 (2011).
- ⁵⁶H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
- ⁵⁷T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- ⁵⁸B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, *J. Comput. Chem.* **18**, 1463 (1997).
- ⁵⁹G. Bussi, D. Donadio, and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).
- ⁶⁰J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
- ⁶¹W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).
- ⁶²W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graphics* **14**, 33 (1996).
- ⁶³D. Eisenberg, R. M. Weiss, T. C. Terwilliger, and W. Wilcox, *Faraday Symp. Chem. Soc.* **17**, 109 (1982).
- ⁶⁴C. Dalgicdir, C. Globisch, C. Peter, and M. Sayar, *PLoS Comput. Biol.* **11**, e1004328 (2015).
- ⁶⁵C. Dalgicdir and M. Sayar, *J. Phys. Chem. B* **119**, 15164 (2015).
- ⁶⁶O. Engin, A. Villa, M. Sayar, and B. Hess, *J. Phys. Chem. B* **114**, 11093 (2010).
- ⁶⁷O. Engin and M. Sayar, *J. Phys. Chem. B* **116**, 2198 (2012).
- ⁶⁸D. Eisenberg, R. M. Weiss, and T. C. Terwilliger, *Proc. Natl. Acad. Sci. U. S. A.* **81**, 140 (1984).
- ⁶⁹See supplementary material at <http://dx.doi.org/10.1116/1.4984009> a list of all MD simulations discussed; timeline for SASA, H-bond, and Coulomb energy; DSSP analysis of parallel tetramer, angle-distance correlation plots, and radius of gyration analysis for backbone and phenylalanine side chains in parallel and anti-parallel tetramers; DSSP analysis of dodecamer in bulk water; and DSSP analysis of parallel and antiparallel tetramer at the air/water interface.