ACS Chemical Neuroscience

Research Article

pubs.acs.org/chemneuro

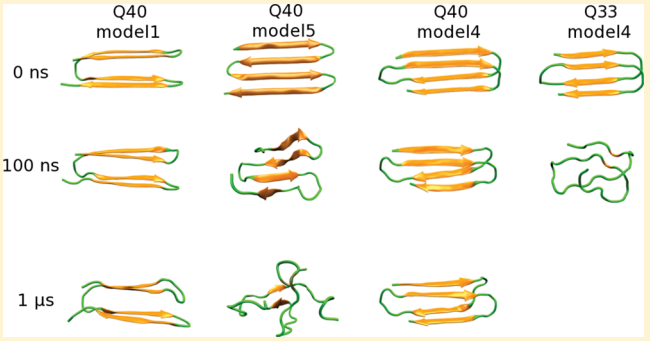
Structural Determinants of Polyglutamine Protofibrils and Crystallites

Viet Hoang Man, Christopher Roland, and Celeste Sagui*

Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202, United States

Supporting Information

ABSTRACT: Nine inherited neurodegenerative diseases are associated with the expansion of the CAG codon. Once the translated polyglutamine expansion becomes longer than ∼36 residues, it triggers the formation of intraneural protein aggregates that often display the signature of cross-β amyloid fibrils. Here, we use fully atomistic molecular dynamics simulations to probe the structural stability and conformational dynamics of both previously proposed and new polyglutamine aggregate models. We test the relative stability of parallel and antiparallel β sheets, and characterize possible steric interfaces between neighboring sheets and the effects of different alignments of the side-chain carboxamide dipoles. Results indicate that (i) different initial oligomer structures converge to crystals consistent with available diffraction data, after undergoing cooperative side-chain rotational transitions and quarter-stagger displacements on a microsecond time scale, (ii) structures previously deemed stable on a hundred nanosecond time scale are unstable over the microsecond time scale, and (iii) conversely, structures previously deemed unstable did not account for the correct side-chain packing and once the correct symmetry is considered the structures become stable for over a microsecond, due to tightly interdigitated side chains, which lock into highly regular polar zippers with inter-side-chain and backbone−side-chain hydrogen bonds. With these insights, we built Q_{40} monomeric models with different combinations of arc and hairpin turns and tested them for stability. The stable monomers were further probed as a function of repeat length. Our results are consistent with the aggregation threshold. These results explain and reconcile previously reported experimental and model discrepancies about polyglutamine aggregate structures.



KEYWORDS: *Aggregation, amyloid fibrils, parallel sheets, antiparallel sheets, polyglutamine diseases, steric zippers*

An expansion of a coding trinucleotide repeat, CAG, translated into an anomalous, expanded polyglutamine (polyQ) tract in the corresponding mutant proteins, is associated with at least nine inherited, fatal neurodegenerative diseases including Huntington's disease (HD), dentatorubralpallidoluysian atrophy (DRPLA), six spinocerebellar ataxias (SCA 1, 2, 3, 6, 7, and 17), and spinal-bulbar muscular atrophy (SBMA).[1] The disease risk and the age-of-onset are correlated with the length of the encoded polyQ expansion, which exhibits a pathogenic threshold of approximately 35−42 glutamine residues, for all but two diseases. Although each disease has a different pathology, they all display polyQ-rich aggregates in neurons once the expansion sequence has crossed the threshold. Thus, it is believed that the neuronal toxicity of polyQ diseases is caused by the polyQ expansion itself, and all the diseases share some common molecular mechanisms of neurodegeneration.[2,3]

Insoluble aggregates of polyQ tracts have been observed *in vitro*, in cell and in animal studies,[4−6] and in diseased human brain tissue.[7] In addition to the mature fibrils, neurons also exhibit a variety of soluble aggregates, sometimes called "protofibrils".[8−10] One possible scenario for the aggregation kinetics of pure polyQ has been described as a nucleation−growth polymerization,[3−5,11−13] but this mechanism can coexist with transient and metastable structures that occur before or simultaneously with fibril formation,[4,5,8−11,14−16] and that may be on- or off-pathway to the formation of aggregates. Although the polyQ-rich aggregates are the most visible cytological features of the polyQ disorders, at present it is not clear whether the formation of aggregates is the cause of toxicity in neuronal cells[17] or whether large aggregates represent a cellular protective response against further toxicity,[15−24] while soluble monomeric or oligomeric intermediates are toxic to the cell.[2,15−26] These findings, common to all amyloid diseases,[24] have spurred efforts to understand the structural attributes of soluble oligomers and amyloidogenic precursors.

The determination of the structure of both soluble oligomers and aggregates of amyloid proteins has motivated extensive, multidecade research efforts. In particular, many structural models have been proposed for polyQ aggregates, such as α-helical coiled coils,[27] β helices,[28] α sheets,[29,30] and β sheets.[31,32]

Received: December 23, 2014
Accepted: January 20, 2015
Published: January 20, 2015

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

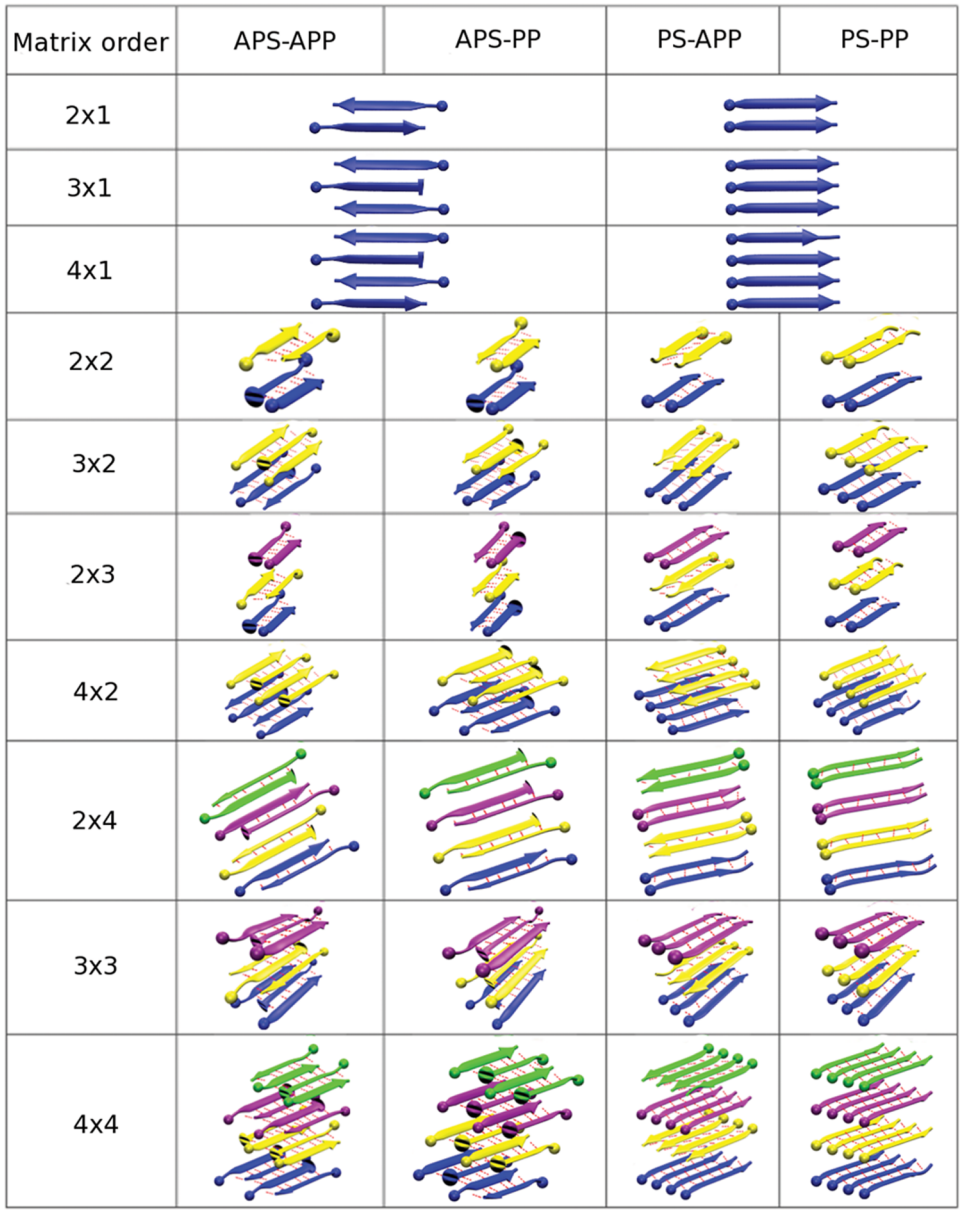


Figure 1. Q_8 protofibril models. APS and PS denote antiparallel and parallel β sheets, while APP and PP denote antiparallel and parallel packing. The balls indicate the N-termini of the peptides.

Current evidence based on polyQ fibers and crystallites supports the presence of cross-β structures,[31−36] common to other amyloid fibrils. In a cross-β structure, the β-strand backbones and side chains are perpendicular to the fibril axis, while the backbone hydrogen bonds are, on average, parallel to the axis. In addition, these structures exhibit dry, "steric-zipper" interfaces[37] between the β sheets (where the residue side chains intermesh with close complementarity).

Given the experimental difficulties to obtain atomic resolution in aggregates, simulations provide invaluable insights into the atomic structure and dynamics of small polyQ aggregates. Unfortunately, approximations used in simulations, although sometimes needed to make progress, can occasionally compromise the biochemical or structural identity of the system and thus inadvertently contribute to the confusion that exists in an already complex field. In our experience, we find that acutely subtle interactions (van der Waals interactions, electrostatic effects, dihedral angles, etc.) that characterize polyQ molecules have not been adequately captured in current coarse-grained models. In particular, it has been shown that polyQ peptides are characterized by long-range correlation effects,[13] that may or may not be present in some of the more simplified models. We believe that the best way to test models is by making contact with high-resolution experiments and providing an explanation to the experimental findings. In the case of polyQ aggregates, a fully atomistic description is needed. An example of this is given by the work of Sikorski and Atkins[32] who in 2005 proposed a model to explain the X-ray diffraction patterns obtained from crystalline fibers and films of a $D_2Q_{15}K_2$ polypeptide[28] (which had originally been interpreted as a hollow, water-filled nanotube[28]). The model successfully explained the main features of the diffraction pattern and, in our estimation, became a turning point in the structural interpretation and understanding of these aggregates. Unfortunately, these aggregates do not lend themselves easily to high- (or medium-) resolution characterization. Thus, such simulations are not very common, although when they do exist,

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

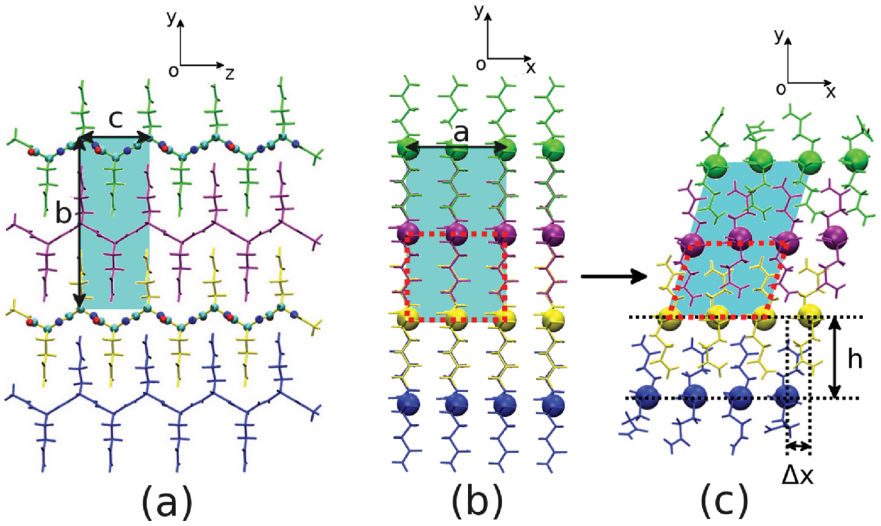


Figure 2. Initial polyQ crystal structure on the *yz* (a) and *xy* (b) planes, and structure on the *xy* plane at $t = 1$ μs (c). The big spheres in panels b and c indicate the peptide axis. Possible unit cells are depicted by colored boxes with lattice parameters $a = 0.950$ nm, $b = 1.660$ nm (cyan) or $b = 0.830$ nm (red dashed line), and $c = 0.657$ nm (parallel) or 0.695 nm (antiparallel).

they can provide invaluable insight into the structure of protofibrils. Another case in point is provided by ref 38, where careful modeling of the system, involving the introduction of various mathematically well-defined crystallographic symmetries, provided very realistic models for Alzheimer's β-amyloid fibrils, in excellent agreement with NMR and X-ray fiber diffraction data.

In this work, we carry out fully atomistic molecular dynamics simulations with explicit solvent and state-of-the-art force fields to probe the structural stability and conformational dynamics of various polyQ aggregates, for a cumulative time of over 23 μs. In particular, we model the oligomeric polyQ aggregates along the lines of ref 32 (a static model that was not tested for stability), but we characterize the possible combinations of peptide stacking and side-chain packing that lead to different steric zipper interfaces. We find that all the models introduced to account for all possible symmetries converge to crystals consistent with available diffraction data, after undergoing spontaneous cooperative side-chain rotational transitions and quarter-stagger displacements on time scales up to microseconds. The fact that such long time scales are required for these conformational transitions also provides a word of caution when interpreting polyQ simulation results. In particular, structures deemed as "polymorphs" could simply be the consequence of inadequate sampling. Moreover, as we discuss below, side-chain packing that lacks the correct symmetry might lead to unstable structures. The structures presented in this work are stable for over 1 μs due to the interdigitated side chains, which lock into highly regular polar zippers with inter-side-chain and backbone−side-chain hydrogen bonds. In addition, we test the stability of aggregated Q_{40} monomers (just above the aggregation threshold) that include different combinations of arc turns or hairpin turns and classify the range of models from unstable to very stable. Finally, we test the stable monomeric models as a function of sequence length and our results are consistent with the aggregation threshold. These results explain and reconcile previously reported experimental and model discrepancies about polyglutamine aggregate structures.

RESULTS

PolyQ expansion is at fault for the formation of intraneuron protein aggregates. Often, these aggregates exhibit patterns attributed to cross-β amyloid fibrils. In this work, we carry out a comprehensive survey of various aggregate models that are compatible with experimental data, linking their relative stability with features of the models (some of which have been proposed before and some of which are new). In addition to studying the stability of parallel and antiparallel β sheets, we map out all the possible steric interfaces that account for the interaction of a sheet with its neighbors, and the effects of different alignments of the side-chain carboxamide dipoles. Once a comprehensive picture for polyQ oligomeric aggregates has emerged, we use this information to build monomeric aggregates of Q_{40} (just above the aggregation threshold) with different combinations of arc and hairpin turns. These models are then tested for stability, and for those that are stable, additional studies are performed to test the sequence-length dependent stability of the models. Our work references at least 75 different structures (either Q_8 oligomers or Q_n monomers, with selected values of n between 29 and 40). Hence, in order to understand the results, we first need to provide a brief description of how the models were built and the nomenclature used to describe them.

Oligomeric Q_8 Structures. We first constructed a dimer structure of the octopeptides Ace-$(Gln)_8$-Nme (Q_8), as described in the Methods section. The initial peptides had fully extended side chains, except for the angle $N-C_\alpha-C_\beta-C_\gamma$, whose initial values are described later. We then used the dimer structure to build $M \times N$ oligomeric "crystals", with unit cell parameters $a = 9.50$ Å (twice the average hydrogen bond distances) in the x direction, side-chain packing distance of 8.33 Å (which results in $b = 8.33$ Å or $b = 16.60$ Å according to the type of packing), and $c = 6.57$ Å (parallel) or 6.95 Å (antiparallel) in the z direction, as sketched in Figures 1 and 2. These unit cell parameters correspond to previous fiber diffraction experiments and models.[32] Here, M and N refer, respectively, to the number of Q_8 peptides in a sheet, and to the number of sheets in the aggregate. In addition to the M and N indexes, each fibril model was characterized by (i) the relative orientation between two neighboring strands in a sheet, which can be a parallel sheet (PS)

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

or an antiparallel sheet (APS), and (ii) the relative orientation between two neighboring stacking sheets, which gives rise to parallel packing (PP) or antiparallel packing (APP). This assumes an alternating pattern between the strands in a sheet and between the sheets in an oligomer, such that every other element has the same orientation in the antiparallel cases (or all the elements have the same orientation in the parallel cases). We considered 34 oligomer models with M in the range 2−4 and N in the range 1−4, as listed in Figure 1. In addition, the initial conditions included peptides whose side-chain carboxamide groups were all oriented in the same direction giving rise to a net total dipole (TD) in the x direction and configurations in which the upper side-chain dipoles canceled the lower side-chain dipoles giving rise to a net zero dipole (ZD) in the x direction. For the packing along the y direction, we generally considered the case where the dipoles of upper and lower side chains in neighboring planes along y have opposite direction. Thus, all in all, we tested 68 oligomer models.

Monomeric Q_{40} protofibrils. Given that the observed β-strand length in a fibril is seven to eight residues,[14,31,39] a Q_{40} peptide can give rise to a maximum of four strands, including the connecting turns and arcs. We thus constructed models of monomeric Q_{40} protofibril by adding arc or hairpin turns (a hairpin joins two strands in the same sheet, while an arc joins two strands in different sheets), as described in the Methods section.

The topological constraint of having a single oriented peptide imposes limitations on the type of four-stranded structure that can be achieved. To characterize the structures, we adopt the following convention. If we approximately divide the long peptide into four strands labeled 1−4 from the amino end to the carboxyl end and denote the connecting loops as arc (A) or hairpin (H), then a notation 1H2A3H4 means that strands 1−2 and 3−4 are connected by a hairpin, while 2−3 are connected by an arc. The possible compact structures that can be built are the following: (i) an APS-APP system with 1−2 and 3−4 β sheets (model 1, 1H2A3H4); (ii) an APS-APP system with 2−3 and 1−4 β sheets (model 2, 1A2H3A4); (iii) an APS-PP system with 1−2 and 3−4 β sheets (model 3, 1H2A3H4); (iv) a PS-APP system with 1−3 and 2−4 β sheets (model 4, 1A2A3A4). As simulations show, the APS-APP and APS-PP Q_8 aggregates used to build the monomer converge to the same structure thanks to a "quarter-stagger" along the x direction. The difference between model 1 and model 3 is that in model 1 the arc connection forces strand 3 to be parallel to 2 and strand 4 parallel to 1, while in model 3 the arc forces strands 3−1 to be parallel and strands 4−2 to be parallel. Notice that under the constraint of a single directed peptide, a PS−PP packing is no longer possible. Of these models, only model 4 is topologically related to a collapsed β helix. Finally, we also consider a planar arrangement (model 5, 1H2H3H4) formed by a 4 × 1 APS system with three hairpins.

In addition, several Q_n monomers with n near the n = 36 threshold were created from the Q_{40} monomers by eliminating residues from the β strands (since hairpins and arcs already have the minimum possible number of residues).

Antiparallel β Sheets Are More Stable than Parallel β Sheets. In this section, we compare the relative stability of parallel and antiparallel β sheets (PS and APS), by evaluating various statistical quantities. A measure of stability is given by the population of residues in β-sheet conformation, with a larger population corresponding to a more ordered and stable structure. In principle, the β-sheet population is useful for all $M \times N$ structures, but we define two further quantities to facilitate the comparison between the single β-sheet models, which include dimers, trimers, and tetramers. First, we define an "L-stranded β sheet", where L varies from 2 to 4, such that at least two additional interpeptide H-bonds must be formed between consecutive β strands in order to pass from a j-stranded β sheet to a $(j + 1)$-stranded β sheet. We then calculate the populations of states that show at least one two-stranded β sheet (P_{bs}^{2s}), one three-stranded β sheet (P_{bs}^{3s}), and one four-stranded β sheet (P_{bs}^{4s}) (a dimer can only show P_{bs}^{2s} population, but a tetramer that shows population in P_{bs}^{4s} will by definition show P_{bs}^{2s} and P_{bs}^{3s} populations). Second, for the particular case of dimers, we classify the structures into S1, S2, S3, S4, and S5 states as shown in Figure 3,

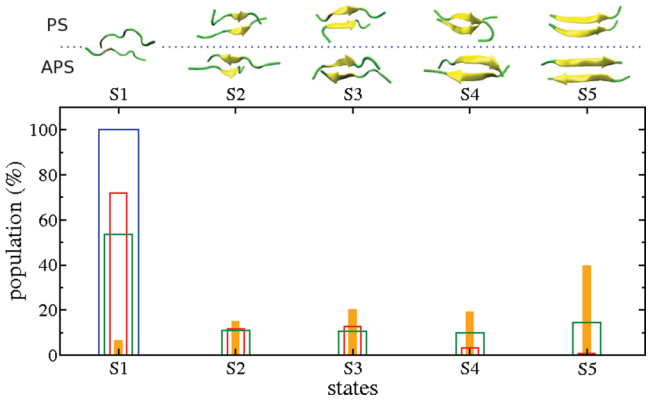


Figure 3. Dimer population of the β-sheet states (S1−S5). S1 state has no residues in β-sheet conformation, S2, S3, and S4 states have two, three, and four residues, respectively, in β-sheet conformation, and S5 states have five or more residues in β-sheet conformation. Orange and red colors are used for ZD and TD side-chain models for the antiparallel dimer, while green and blue colors show ZD and TD side-chain models for the parallel dimer. Results are obtained from 100 ns simulations.

such that S1 states have no residues in β-sheet conformation, S2, S3, and S4 states have two, three, and four residues, respectively, in β-sheet conformation, and S5 states have five or more residues in β-sheet conformation. For all cases, we also compute the "broken time" (τ_b) parameter, which is the first time that the ordered state is broken. Unless otherwise specified, the results shown in this section are for the initial configurations with zero side-chain dipoles (ZD models), but similar trends are observed for the total dipole (TD) models.

Single β-sheet models. Table S1, Supporting Information, shows the L-stranded populations for the single β-sheet models, along with the broken time τ_b. The increased stability of the antiparallel sheets is reflected by the higher population and higher τ_b for all structures. For the parallel dimers, P_{bs}^{2s} = 29% and τ_b = 16 ns, while for the antiparallel dimers, P_{bs}^{2s} = 93% and τ_b = 60 ns. Notice that the dimer necessarily recombines after it breaks the first time (otherwise in 100 ns simulations, the populations would just be 16% and 60%, respectively).

Figure 3 shows the dimer distribution along the five states during the 100 ns of simulation. From this picture, it is clear that the TD models (red for APS and blue for PS) are not stable. The ZD models are more stable, with the ZD APS model (orange) considerably more stable than the ZD PS model (green). Clearly, the antiparallel sheet is more stable than the parallel sheet, without exceptions. Supporting Information Figure S1 and Figure 4A show the population for the number of residues in β-sheet conformation for the trimers and tetramers for the ZD model of side chains. The 4 × 1 APS system is the most stable, pointing to cooperative effects.

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

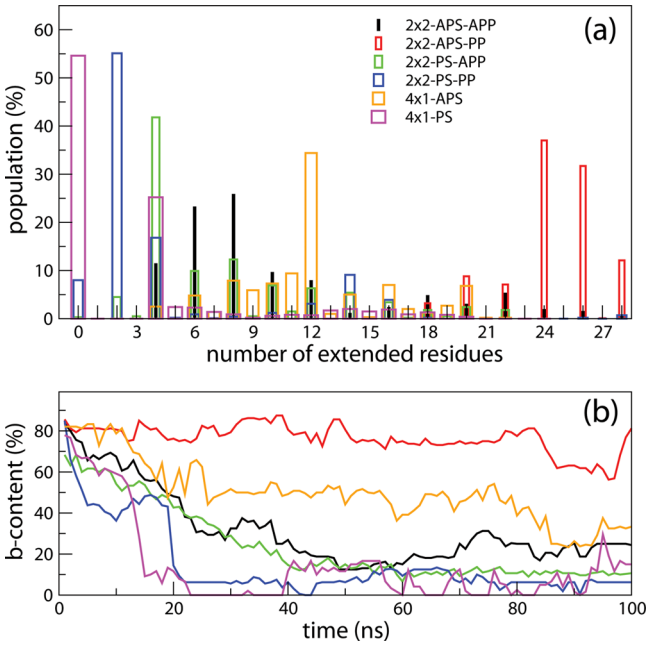


Figure 4. Population of states as a function of the number of extended residues (a) and time dependence of the β content (b) of the 2 × 2 and 4 × 1 Q_8 tetramers. Results are obtained from 100 ns of MD simulations.

Multiple β-Sheet Models. For the $M \times N$ systems, with $N > 1$, in addition to the relative order of the strands within a hydrogen-bonded sheet (PS or APS), we have to consider the stacking of the sheets, which gives rise to parallel packing (PP) when the strands on top of each other point in the same direction or antiparallel packing (APP) when the strands on top of each other point in the opposite direction. This gives rise to four possible combinations, as listed in Figure 1, each with one of two initial side-chain dipoles, either zero (ZD) or total dipole (TD) in the xy plane. Table 1 lists for the ZD case the broken times τ_b and populations of the configuration with the highest fibril order. For an $M \times N$ system, the highest fibril order is represented by N sheets, each of which is M-stranded. For instance, for a 3 × 2 system, the highest fibril order occurs when each of the two sheets is a three-stranded β sheet, as defined before.

We first compare the stability of the 2 × 2 models. Figure 4 shows that the APS-PP system is mainly distributed in β-rich states, while the PS−PP system mainly populates β-poor states. The two APP systems, APS-APP and PS-APP, have intermediate values of population, in agreement with Table 1. The ordering in terms of decreasing stability is APS-PP, APS-APP, PS-APP, and PS−PP, and the APS systems are always more stable than the PS systems, as also shown in Supporting Information Figures S2 and S3 for hexamers and octamers. The stability of all systems increases with the order of the matrix; the 4 × 4 models are still stable at 1000 ns.

Impact of Matrix Order on Stability. In this section, we present results that show the impact of matrix order on stability. We define the matrix order $M \times N$ as a matrix arrangement, without considering the directional relation between the peptides. Here, we describe the effect of the matrix order on the stability of oligomers containing the same number of Q_8 peptides. Thus, the tetramers include the 4 × 1 and 2 × 2 matrices, the hexamers include 3 × 2 and 2 × 3 matrices, and the octamers include 4 × 2 and 2 × 4 matrices.

The most ordered oligomer in a 4 × 1 system has $P_{bs}^{4s} = 26\%$ and $\tau_b = 18$ ns for APS, and $P_{bs}^{4s} = 11\%$ and $\tau_b = 9$ ns for PS, while the most ordered oligomer in a 2 × 2 system corresponds to APS-PP, with $P = 100\%$ and $\tau_b > 100$ ns, as shown in Table 1. This table shows that both possible packings of APS and PS sheets have populations in the 2 × 2 arrangement higher than the P_{bs}^{4s} populations in the 4 × 1 system, indicating that a higher stability is associated with the 2 × 2 system.

Similar considerations were applied to the hexamers and octamers. The 3 × 2 and 4 × 2 models are more stable than 2 × 3 and 2 × 4 ones, respectively (Supporting Information Figure S2 and Figure S3). These results can be explained in terms of the compactness of the structure quantified by the aspect ratio. Since all the structures have the same size along the z axis (the length of the peptide), we simply use the x and y axis to define the aspect ratio as the ratio between the longest and shortest side for the initial structures, as presented in Table 1. For oligomers composed by the same number of peptides, the smaller aspect ratio is indicative of a more compact structure, as it is visually obvious in Figure 1. The more compact structures maximize the number of inter- and intrapeptide hydrogen bonds and the van der Waals contacts between side chains, and therefore the stability of the aggregate. Table 1 also lists the maximum number of *possible* hydrogen bonds for the structures. Supporting Information Figure S4 shows the maximum number of hydrogen bonds for different aggregates. For every peptide, two middle residues are used to calculate the total number of possible hydrogen bonds. For instance, there are 12 possible hydrogen bonds for the 4 × 1 system, and 14 hydrogen bonds for the 2 × 2 system.

Crystal Structure and Side-Chain Packing. Due to the long stability of all the 4 × 4 models, we extended the simulation times of these systems up to 1 μs or longer. These systems

Table 1. Compactness and Stability of Multiple β-Sheet Models[a]

system		AR	PHBonds	APS-APP		APS-PP		PS-APP		PS−PP	
				P	τ_b	P	τ_b	P	τ_b	P	τ_b
4-mer	4 × 1	3.43	12 (1.500)	26	18	26	18	11	9	11	9
	2 × 2	1.75	14 (1.750)	38	32	100	>100	30	25	13	10
6-mer	2 × 3	3.49	24 (2.000)	100	>100	100	>100	75	39	20	17
	3 × 2	1.14	26 (2.167)	100	>100	100	>100	100	>100	74	63
8-mer	2 × 4	5.24	34 (2.125)	100	>100	100	>100	80	61	44	31
	4 × 2	1.71	38 (2.375)	100	>100	100	>100	100	>100	100	>100

[a]Shown are the aspect ratio (AR) for the initial structures, the maximum number of possible hydrogen bonds (PHbonds), population P (in %), and broken time τ_b (in ns) for the configuration with the highest fibril order (which contains N M-stranded β sheets, as defined in the text). PHbonds are counted by selecting two middle residues for every peptide in the system (see Figure S4 in Supporting Information) Numbers in parentheses show the PHbonds per inner residue. Results are obtained from 100 ns simulations.

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

contain the 2 × 2 subsystems in the center, which are then surrounded by protein and can therefore be considered as a basic unit of an aggregate structure.

PolyQ Cross-β Structure. Here, we briefly review the known characteristics of polyQ cross-β structures. In protein β sheets, β strands (the polypeptide chains) are connected laterally by at least two hydrogen bonds: if $C_{\alpha,i}$ and $C_{\alpha,j}$ face each other in two adjacent strands, then residue i may form hydrogen bonds to residues $j - 1$ and $j + 1$. Each strand exhibits 2-fold helical conformation (2_1) with a distance between two residues along the backbone axis (the crystallographic c axis) of 3.5 Å. Neighboring strands are aligned so that their C_α are also aligned with the side chains pointing in the same direction: successive side chains point up or down, perpendicularly to the plane of the sheet, thus reflecting the 2_1 symmetry, which is a direct consequence of the sp^3 tetrahedral chemical bonds of C_α and gives the sheet its typical pleated appearance. Thus, the crystallographic c axis is ∼7.0 Å. The a axis lying in the average hydrogen bond direction is orthogonal to the c axis with a typical intrasheet distance between adjacent strands of ∼4.75 Å. Crystallographic periodicity demands twice this distance in APS sheets, that is, $a \approx 9.50$ Å. The side chains are along the b axis and orthogonal to the ac plane on alternating faces of the sheet, with intersheet stacking distances that depend on the nature of the amino acid. For polyQ, a distance between α carbons in adjacent layers of ∼8.30 Å has been observed experimentally.[28,31,32,40] There is an intrinsic tendency for neighboring sheets to slip with respect to each other by around 0.47/2 in the a direction.[32,41] A cross-β structure is formed when the c axis along the backbone atoms is orthogonal to the axis of the crystal, which extends along the a axis, that is, parallel to the backbone hydrogen bonds. In particular, amyloids display characteristic fiber diffraction patterns that are attributed to a cross-β structure.[42,43]

Initial Structures. As described in the Methods section, we modified the initial, fully extended (with 180° torsion angles) Q_8 peptides such that the backbone torsion angles acquired typical values for parallel and antiparallel sheets and then rotated the side chains so that these lay along the y axis (or initial b axis) perpendicular to the ac plane of the sheet. All other torsion angles were left at 180°.

Considering that a zipper-like mechanism is supported by both experimental data and models, we proceeded to model the initial structures with an intrasheet distance along the hydrogen bonds of 4.75 Å and an intersheet distance of 8.3 Å. We noticed also that there are several ways of packing the side chains along a (generally by rotating the $N-C_\alpha-C_\beta-C_\gamma$ angle) so that there is contiguous hydrogen bonding between the carboxamide groups. Figure 5 shows possible side-chain packing along a. The parallel sheet in Figure 5a shows a perfect string of hydrogen bonds that form when the final Q_8 peptides are brought together as rigid objects. This configuration also satisfies the 2_1 symmetry along c of β strands. Instead, when alternating strands are flipped (as a rigid object) in order to form an antiparallel sheet, the side chains would experience strong electrostatic clashes between the side-chain carboxamide groups (Figure 5b). Thus, in the APS case, it is necessary to further rotate alternating side chains. This can be achieved by rotating the torsion angle $N-C_\alpha-C_\beta-C_\gamma$ of alternating peptides by 180°, as shown in Figure 5c. In this and other figures, purple circles with a + sign indicate that the $N-C_\alpha-C_\beta-C_\gamma$ torsion angle is positive (with an initial value around 75°), and a − sign within the purple circles indicates that the $N-C_\alpha-C_\beta-C_\gamma$ torsion angle has been rotated by 180° (with an initial value around −105°). Alternatively, one could rotate a torsion angle further along the chain, such as the $C_\beta-C_\gamma-C_\delta-N_\varepsilon$ torsion shown in Figure 5d, where all the $N-C_\alpha-C_\beta-C_\gamma$ torsions are positive (we have not done this; Figure 5d is instead a final configuration corresponding to Figure 5c). The strands in the initial conditions, Figure 5a–c, all have the 2_1 symmetry and a net zero dipole (ZD). In addition, it is possible to rotate the side chains on one side of the sheet plane, such as shown for the PS and APS sheets in Figure 5e,f. This breaks the 2_1 symmetry for the strand and results in a net side-chain dipole on the ab plane (because the polar hydrogen bonds on both sides of the β sheet are pointing in the same direction). We call this model the total dipole (TD) model.

Our simulations used as initial conditions the ZD (Figure 5a) and the TD (Figure 5e) for the PS systems and the ZD (Figure 5c) and the TD (Figure 5f) for the APS system. Figure 5d was not

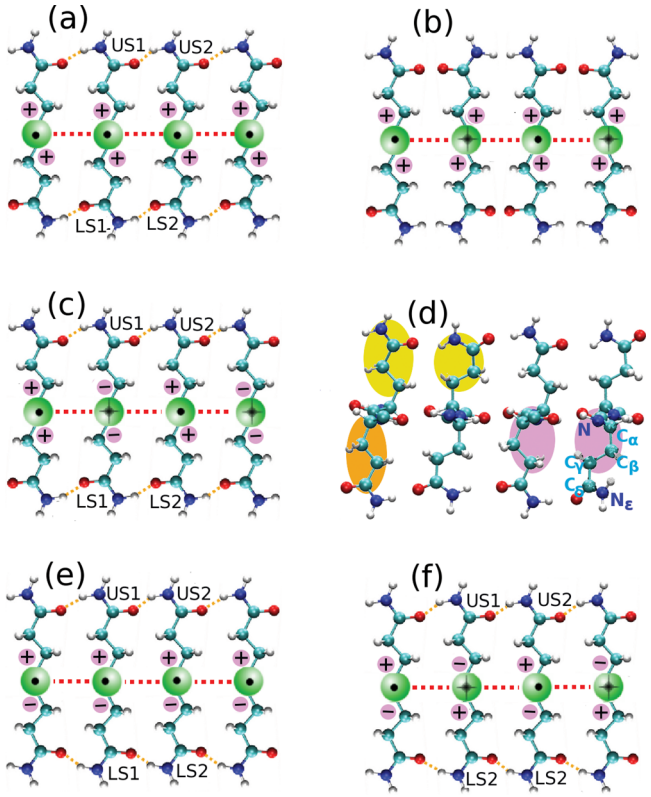


Figure 5. Different possible side-chain packing arrangements. Oligomers are shown in the xy plane, with the peptide backbone chain illustrated by a green sphere with a black circle or a plus mark for a chain coming out of or into the plane. N, O, H, and C atoms are shown in blue, red, white, and cyan spheres, respectively. Orange dashed lines show hydrogen bonds between side chains, while red ones show hydrogen bonds between backbone atoms. (a) PS showing 2_1 symmetry of a β strand and ZD on the xy plane. (b) APS with "rigid" side chains unable to form hydrogen bonds. (c) APS derived from the structure in panel b by rotating the side chains of the 2nd and 4th peptides around the $N-C_\alpha-C_\beta-C_\gamma$ torsion angle by 180° in order to allow side-chain hydrogen bonds; it has β strands with 2_1 symmetry and ZD on the xy plane. (d) APS showing the $N-C_\alpha-C_\beta-C_\gamma$, $C_\alpha-C_\beta-C_\gamma-C_\delta$ and $C_\beta-C_\gamma-C_\delta-N_\varepsilon$ torsion angles in purple, orange, and yellow colored ellipses, respectively. The configuration displays 2_1 symmetry and ZD on the plane, and all the $N-C_\alpha-C_\beta-C_\gamma$ torsions are positive. (e) PS with $N-C_\alpha-C_\beta-C_\gamma$ torsion angle of lower side chains of every peptide flipped by 180°, exhibiting no 2_1 symmetry and a TD on the xy plane. (f) APS with no 2_1 symmetry and a TD, similar to panel e.

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632–645

an initial condition, but (one layer of) the resulting structure of a 4 × 4 APS system that evolved from Figure 5c to Figure 5d.

Time Evolution and Crystal Structure. Notice that even though we set the initial conditions according to a previous model that is consistent with diffraction data, we do not "overconstrain" the initial conditions. Thus, for instance, the initial side chains are fully extended, and we do not implement the experimentally observed "quarter-stagger".[41] This is important: the models are not just stable or unstable in the sense of keeping close to their initial conditions but actually evolve in time: in particular, the stable systems evolve toward more stable structures, which are not quite the same as those used in the initial conditions but whose typical crystal parameters are still consistent with the available experimental data.

The time evolution of the ZD systems shows spontaneous quarter-stagger (where adjacent sheets slip with respect to each other by $a/4$ in the a direction). Figure 2 shows initial configurations the yz plane (a) and xy plane (b) and the configuration at 1 μs on the xy plane (c) for a ZD 4 × 4 APS-PP system. The original fiber diffraction pattern[28,32] seems to be sensitive only to the positions of the α-carbons, showing sharp equatorial diffraction signals for 4.75 and 8.33 Å and a broad meridional arc of 3.60 Å. All eight 4 × 4 lattices at the end of the simulation share these values, with small deviations. Thus, one can formally attribute a monoclinic lattice with $c \simeq 2 \times 3.60$ Å for all lattices, $a \simeq 4.75$ Å for the PS systems, and $a \simeq 9.50$ Å for the APS systems. The intersheet stacking is also around 8.33 Å for all systems, but whether one takes that value or twice that value for the axis b depends on the initial stacking of the side chains. In the initial conditions for the TD APS and PS systems (Figure 5e,f) and the ZD APS systems (Figure 5c), the side-chain rotations can be as shown, or the "+" and "−" side chains can be exchanged resulting in zippers with dipoles in the opposite direction. This results in two possible ways of stacking the systems along y. Thus, ZD PS−PP has $b \approx 8.3$ Å, and the other lattices have $b \approx 8.3$ Å or $b \approx 16.6$ Å, according to the relative directions of lower side-chain dipoles in one sheet and the upper side-chain dipoles in the adjacent sheet along y, combined with whether they are ZD or TD in the xy plane. Table 2 gives the crystal parameters averaged over the last 700 ns of the simulation.

Table 2. Average Crystal Parameters for the 4 × 4 ZD Systems[a]

model 4 × 4	a	b	c	Δx
APS-APP	9.68 ± 0.05	8.25 ± 0.10	7.15 ± 0.02	2.42 ± 0.16
APS-PP	9.68 ± 0.05	8.25 ± 0.11	7.14 ± 0.02	2.41 ± 0.19
PS-APP	9.94 ± 0.03	8.34 ± 0.09	7.19 ± 0.05	2.41 ± 0.12
PS−PP	9.97 ± 0.03	8.14 ± 0.08	7.22 ± 0.05	1.96 ± 0.11

[a]The parameters (in Å) are shown in Figure 2, and they are obtained as an average over the last 700 ns of the 1 μs simulations. Here, b refers to the distance between the closest α carbons in two adjacent sheets, and Δx is the displacement along x between adjacent β-sheet layers.

Supporting Information Figure S5 shows the time evolution of the crystal parameters for the ZD models. The quarter-stagger takes place fairly fast in the APS-APP model, and it takes about 220 ns in the APS-PP model. After this transition, the two systems have converged to the same structure (Figure 2), with the same crystal parameters (Table 2). With respect to the PS systems, while the PS-APP system seems to have almost completed the quarter-stagger displacement in 1 μs, the PS−PP system has not undergone a complete quarter-stagger displacement yet. The crystal parameters, especially those for the APS systems, are consistent with general features observed experimentally, as described before.

Next, we consider the evolution of the ZD side chains. Figure 6 and Supporting Information Figure S6 show results for the APS-PP and APS-APP systems. The figures display various torsion angles of the "upper side chains" 1 and 2 (US1 and US2) and the "lower side chains" 1 and 2 (LS1 and LS2) of two middle neighbor peptides, as shown in Figure 5 (where the side chains of only the four inner residues are labeled). The evidence indicates that the rotation of the N−C_α−C_β−C_γ angles by 180° is not favored. The ZD APS systems have as the initial configuration in the ab plane that shown in Figure 5c: the side chains of peptide 1 have been flipped, so that the N−C_α−C_β−C_γ angles of US1 and LS1 have negative values. These, however, flip back, as shown in the upper panels of Figure 6 and Supporting Information Figure S6, and become centered around the glutamine equilibrium value of 66°. The rotation of the N−C_α−C_β−C_γ angle of US1 and LS1 changes the value of the C_β−C_γ−C_δ−N_ε angle (bottom panels in Figure 6 and Supporting Information Figure S6), which changes from 180° to 0°, thus preserving the polar zipper mechanism as shown in the final configuration presented in Figure 5d. The flipping of the N−C_α−C_β−C_γ and C_β−C_γ−C_δ−N_ε angles is coupled to the Δx displacement, as shown in the bottom panel of Supporting Information Figure S5, which results in the quarter-stagger conformation. By contrast, the ZD PS systems, whose initial configuration in the ab plane is shown in Figure 5a, have initial angles naturally favored by the amino acid, and thus the side chains do not experience (long-lived) flipping. The PS-APP and PS-PP systems are shown in Supporting Information Figures S7 and S8, respectively.

The case of the TD side chains is slightly more complicated. The TD APS systems have Figure 5f as initial configuration in the ab plane: the initial flipped chains are US1 and LS2, in alternating peptides. At the end of the 1 μs simulations, we found that these chains had not flipped to their equilibrium value around 66°, but there were clear signs of instability, as can be seen in Supporting Information Figure S9 for the APS-APP system. Thus, we ran this system for an additional 0.5 μs during which we can observe complete flipping of both the US1 and LS2 side chains, coupled with some large oscillations of US2. The flipping of the chains is synchronous to the settling of the system in the quarter-stagger configuration. Since the time scale for these transitions are quite large (and therefore computationally expensive), we took the other three TD systems at the end of the 1 μs simulations, slightly displaced the planes along the a direction in order to complete the incipient "quarter-stagger", and restarted the simulations allowing the side chains to relax on their own. We see that the all torsions associated with US1 and LS2 quickly relax, the N−C_α−C_β−C_γ angles center at 60°−65°, and their C_β−C_γ−C_δ−N_ε angles flip to 0°. Supporting Information Figures S10−S12 show the last 300 ns that followed the 1 μs "intervention".

Figure 7 shows final configurations for the 4 × 4 systems. The figure shows the four innermost peptides on the yz plane for an APS-PP crystal (Figure 7b), and four inner residues within those peptides for each of the possible packing arrangements on the xy plane (Figure 7a), displaying details of the side-chain packing. Notice that the N−C_α−C_β−C_γ flipping preserves the dipolar character of the side chains in the ab plane, ZD or TD. The quarter-stagger creates an additional set of hydrogen bonds.

Finally, Figure 8 shows the hydrogen bond distances for the ZD 4 × 4 systems. The two top panels show the distance between

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

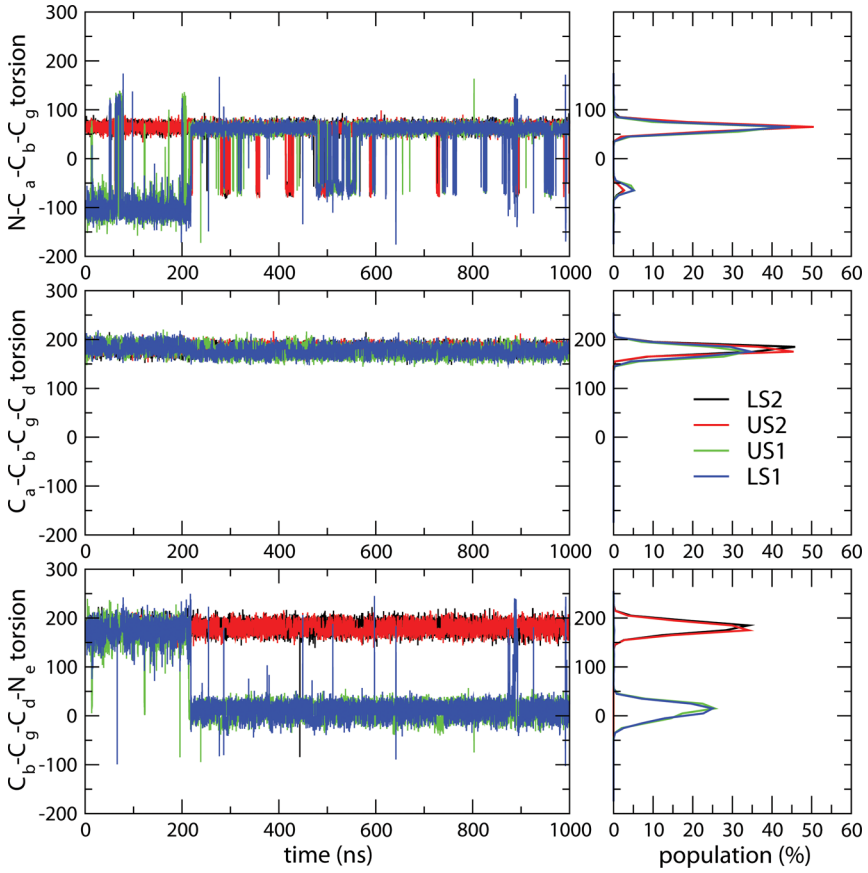


Figure 6. Time dependence and distribution of torsion angles of central residues in a 4 × 4 APS-PP model with ZD side chains. The initial and final conformations on the *ab* plane are given in Figure 5c,d. The distribution was calculated from the last 700 ns of simulation time.

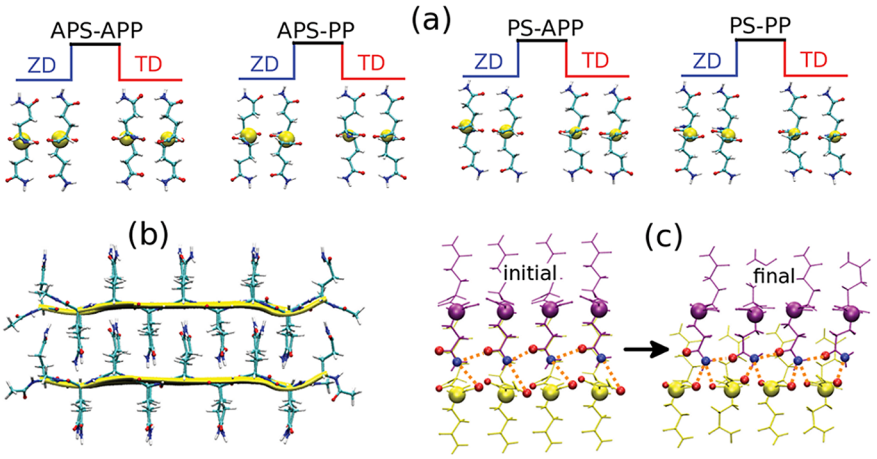


Figure 7. Final side-chain configurations for the 4 × 4 aggregates. For ease of visualization, four inner peptides in the 4 × 4 aggregates are selected, as shown for an APS-PP system on the *yz* plane in panel b. Within these peptides, four inner residues are selected and the configurations of their side chains are depicted on the *xy* plane, as shown in panel a for different arrangements and for both the ZD and TD cases. In panel c, a cut through two sheets (represented by yellow and purple) of an APS-PP ZD system is shown on the *xy* plane at the initial ($t = 0$) and final times ($t = 1$ μs). Red and blue balls represent O and N atoms. The bigger (smaller) red balls indicate O atoms closer to (further from) the viewer. Hydrogen bonds linking side chains with each other and with backbone atoms are shown in orange for the selected group of atoms (main-chain hydrogen bonds are omitted).

the O_ε and N_ε atoms in the US1−US2 and LS1−LS2 side chains. The two lower panels show the distance between the side-chain N_ε of side chains US1 and LS1 and the main chain O atom of the nearest sheet. In this figure, the distances to only one of the main chain O atoms is shown, but after the initial relaxation, the N_ε is equidistant to the two main-chain O atoms in the nearest sheet, as shown schematically in Figure 7. This is a consequence of the quarter-stagger conformation. This shows that the polar zipper mechanism is sustained by a stable network of hydrogen bonds, with main-chain−main-chain hydrogen bonds, side-chain−side-chain hydrogen bonds, and side-chain−main-chain hydrogen bonds.

Stability of the Q_n Monomeric Forms. Figure 9 shows the initial and various time snapshots in the evolution of the five

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

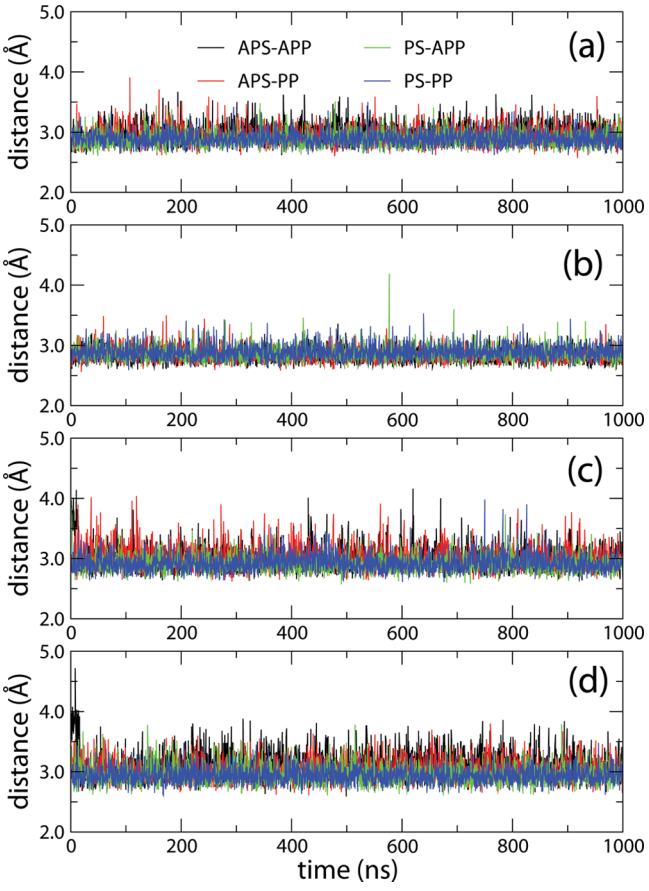


Figure 8. Hydrogen bonding interaction between the side-chain O_ε and N_ε atoms (a and b) and between the side-chain N_ε and the main chain O atoms (c and d). Panels a and b show the side-chain hydrogen bond distances of US1−US2 residues and LS1−LS2 residues, respectively. Panels c and d show the hydrogen bond distances between US1 and LS1 side chains and the main chain of the nearest sheet.

models of Q_{40} aggregates, while Figure 10 illustrates their stability as the number of residues is changed around their threshold. Models 1 and 3 are "β sheet stacks". Both models contain two hairpins and one arc turn (1H2A3H4), the main difference being that the arc in model 1 connects two strands on the same side of both sheets (thus we call it "lateral β sheet stack") while the arc in model 3 connects strands on the opposite side of the sheets ("diagonal β sheet stack"). Due to the quarter-stagger, the "side" and the "diagonal" in this arrangement are the same length of ∼8.3 Å, so the arcs are of the same length. Model 2, with two arcs and one hairpin (1A2H3A4) will be referred to as a "β double hook". Model 4 contains three arcs (1A2A3A4) and can be interpreted as a collapsed β helix. Finally, model 5 (1H2H3H4) is simply a single-sheet 4 × 1 β structure connected by three hairpins.

Just a glance at Figure 9 is enough to see that models 1−4 are stable in the 1 μs time scale, while model 5 is unstable. Figure 11 shows the root-mean-square deviation and the percentage of β content as a function of time. The percentage of β content is computed with respect to the initial β content of each aggregate, because this differs for the different models (with model 5 having the maximum number of residues, 37, in β-sheet conformation at time zero, and model 4 the minimum, 31). This confirms that models 1, 2, and 4 are perhaps the most stable, model 3 is less stable, and model 5 is unstable.

In addition, it is interesting to see how the stability of these models varies as the numbers of Q residues oscillates around the threshold of 35−36. Figure 10 shows that model 1 is very stable: there is apparently no difference between 40 and 37 residues and even 33 residues seem to be stable at t = 500 ns. However, Q_{29} is indeed too short, and the aggregate is already unfolded at t = 50 ns. At t = 500 ns, model 2 is stable for 35 residues and becomes unstable for 31 residues, and model 4 is stable for 37 residues but becomes quickly unstable for 33 residues.

DISCUSSION

In each of the following subsections, the first paragraphs summarize one or more results from this study, while the subsequent paragraphs present a brief comparison to other work.

Impact of Matrix Order, Relative Orientation of β Strands, and Dipolar Order in Q_8 Aggregates. One result of this study is that while the stability of an oligomer generally increases with the number of monomers, the way these monomers are assembled is also very important. For the same number of monomers, compact structures are more stable than loose structures (as characterized, for instance, by aspect ratios or number of possible hydrogen bonds). This is a consequence of the cooperative character of the polyQ assembly. Thus, for octamers, the more compact 4 × 2 (two sheets of four strands each) aggregates are more stable than the 2 × 4 aggregates (four sheets of two strands each), which exhibit larger aspect ratios. We expect this trend to apply only for a small number of monomers, after which a "saturation point" kicks in, and the corresponding fibrils or crystals can grow with very large aspect ratios. For both single β-sheet models and multiple β-sheet models up to octamers, antiparallel β sheets are more stable than parallel β sheets. On the other hand, both APS and PS conformations in the larger, very compact 4 × 4 aggregates are stable for longer than 1 μs. In the loose structures, conformations with a total side-chain dipole on the *ab* plane are less stable than those with a zero side-chain dipole. The side-chain dipole is due to the same orientation of the side-chain carboxamide groups both above and below the sheet plane (Figure 5e,f). This effect is stronger in the single β sheets. As the structures are packed along the *y* direction, this destabilizing effect becomes less pronounced. In the 4 × 4 aggregates, both TD and ZD aggregates are stable.

In amyloid fibrils, different peptide sequences favor different β-sheet structures. For instance, Aβ peptides commonly form parallel β sheets in fibrils, while antiparallel ones only occur in aggregates formed by relatively short peptides.[44] Parallel β sheets were observed in Q-rich (with other amino acids) hexapeptides.[37] On the other hand, ThT fluorescence assay, X-ray diffraction, solid-state NMR spectroscopy, and infrared spectroscopy studies suggest that polyQ peptides form antiparallel β sheets within fibrils.[14,31,34,35] In fully atomistic simulations, the total energy of Q_8 dimers in implicit solvent was measured to be lower for antiparallel than parallel β sheets,[30] and MD simulations of various short polyQ dimers also in implicit solvent showed stability for antiparallel β sheets for over 50 ns but not for parallel ones.[45] In explicit water, 10 × 2 Q_6 and 4 × 2 Q_{15} aggregate models (with a total of 120 glutamines) based on the crystal structure of the peptide NNQQNY were found to be stable for both parallel and antiparallel conformations.[46] However, these peptides were capped with four charged lysine and glutamate amino acids, which enhance stability, and simulations were run only up to 20 ns. For comparison, our smaller 4 × 2 Q_8 PS aggregates (64 glutamines) only show signs of instability around 100 ns.

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

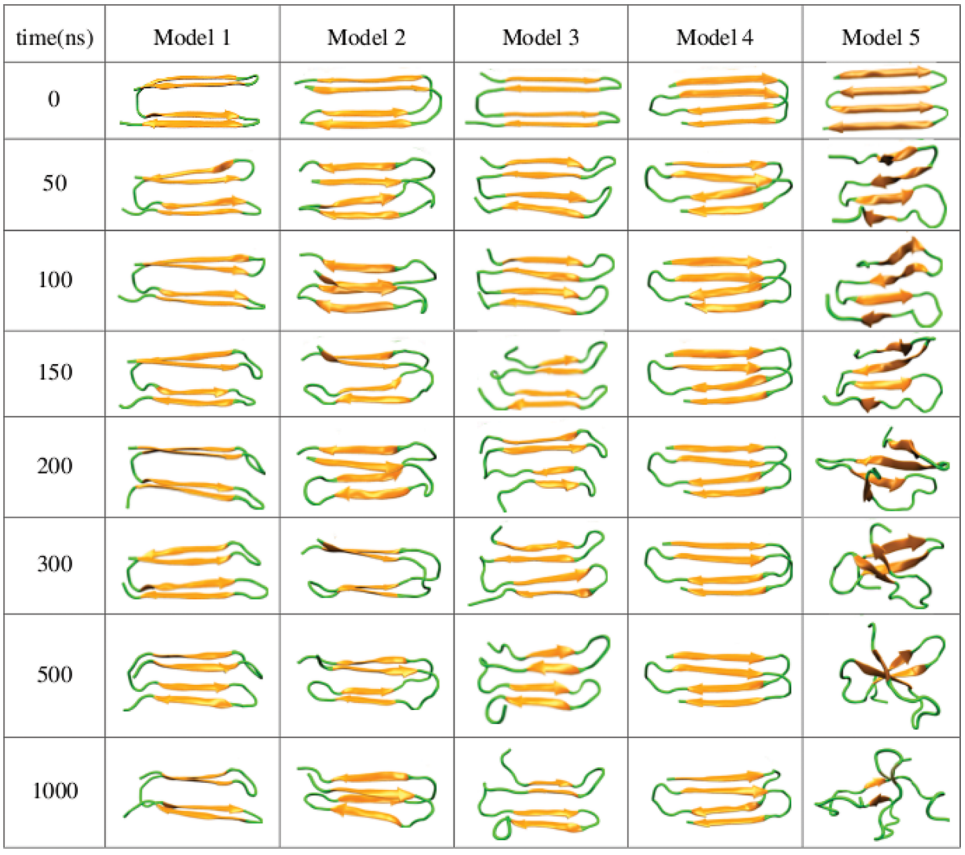


Figure 9. Time evolution of the Q_{40} fibril models. β strands are shown in orange color, while coil, arc, and hairpin structures are shown in green.

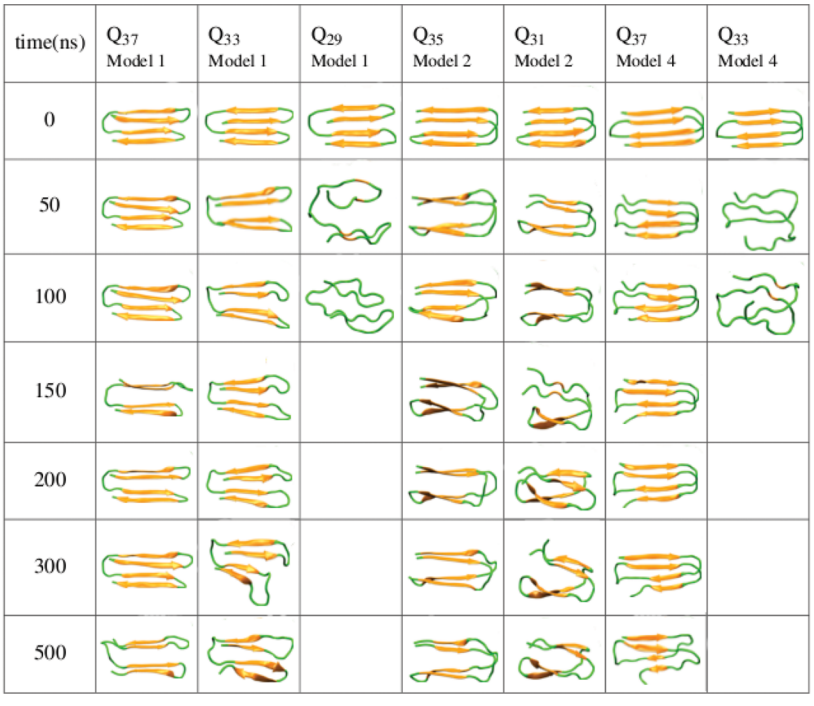


Figure 10. Time evolution of monomeric fibril models of various lengths. β strands are shown in orange color, while coil, arc, and hairpin structures are shown in green.

Time Evolution, Steric Zipper, And Crystal Structure in the 4 × 4 Q_8 Aggregates. We started the simulations with four different initial conditions (Figure 5a,c,e,f) with either ZD or TD on the *ab* plane (combined with two packings, APP or PP, along the *y* direction). In these initial structures, the side chain was considered a fully extended rigid chain, except that the angle N−C_α−C_β−C_γ was rotated 180° in three of the four cases, as suggested by a successful model,[32] in order to prevent clashing of

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

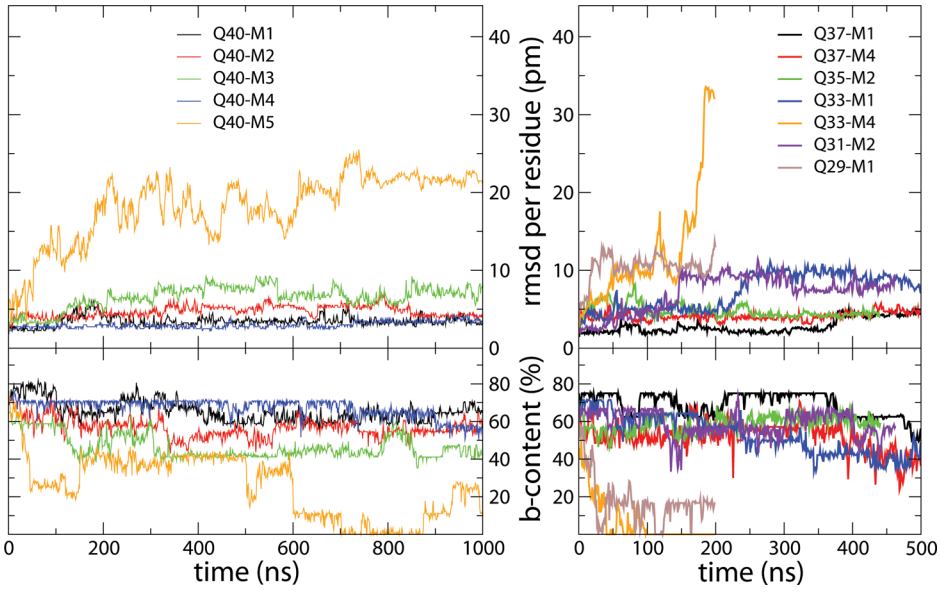


Figure 11. Time dependence of the backbone root-mean-square deviation (rmsd) with respect to the initial conformer and time evolution of the β content of each monomer. The right panels show data for the five Q_{40} models, while the left ones present data for other repeat lengths varying from 37 to 29 residues. Labels M1, ..., M5 refer to model 1, ..., model 5.

polar groups. The side chains were interdigitated[32,37] (Figure 2) but without "quarter-stagger". We then let the various systems (APS/PS, APP/PP, TD/ZD) evolve for 1 μs or longer. The time evolution of the system showed several interesting results: (i) The steric zipper with interdigitated side chains was not only preserved during the 1 μs simulations but also reinforced by the quarter-stagger displacement, as shown in Figure 7. The polar zipper was strongly held together by interbackbone amide hydrogen bonds, inter-side-chain amide hydrogen bonds, and hydrogen bonds between the side-chain N_ε and the two main-chain O atoms in the nearest sheet. (ii) The N$-C_\alpha-C_\beta-C_\gamma$ rotation of 180° was not favored; in all but two cases, the chains flipped back to the Q favored value of approximately 66° (Figure 6 and Supporting Information Figures S5−S12; in the PS ZD cases, the initial torsions already have the preferred value). This rotation changed the values of the $C_\beta-C_\gamma-C_\delta-N_\varepsilon$ angles so that the side-chain hydrogen bond pattern and the dipolar character of the layers were preserved. (iii) The quarter-stagger motion and the flipping of the side chains took place in different, albeit long, time scales, such as 220 ns (APS-PP, ZD) and 1 μs (APS-APP, TD). (iv) After these transitions, all the structures converged to a monoclinic crystal structure (Figure 2) whose lattice parameters (Table 2) are consistent with experimental observations.[28,31,40]

The resulting side-chain conformations after the N$-C_\alpha-C_\beta-C_\gamma$ rotation can be seen in Figure 7. In this figure, four inner peptides in the 4 × 4 aggregates are selected, as shown for an APS-PP system on the *yz* plane in panel b. Within these peptides, four inner residues are selected, and the configurations of their side chains are depicted on the *xy* plane, as shown in panel a. Clearly, the flipping of the N$-C_\alpha-C_\beta-C_\gamma$ angle preserves the dipolar character of the side chains in the *ab* plane, ZD or TD. In panel c, a cut through two inner sheets of an APS-PP ZD system is shown on the *xy* plane at the initial ($t = 0$) and final times ($t = 1$ μs). Essentially, the final configuration differs from the initial configuration in the flipping of the N$-C_\alpha-C_\beta-C_\gamma$ angles and subsequent relaxation of the side chains (helped by the quarter-stagger), which results in an additional set of hydrogen bonds connecting the side chains to the backbones.

Experimentally, it has been known for many decades that Q-rich peptides exhibit hydrogen bonding not only for the backbone but also for the carboxamide groups in the side chains.[40] However, a detailed atomic model remained elusive for many years, until 1994 when Perutz et al. proposed the "polar zipper model" (PZ model).[47] Later, Sikorski and Atkins[32] proposed a related model to explain the X-ray diffraction patterns obtained from crystalline fibers and films of a $D_2Q_{15}K_2$ polypeptide[28] (which had originally been interpreted as a hollow, water-filled nanotube[28]). The model successfully explained the main features of the diffraction pattern but was not tested for stability. We will refer to this model as the SA model. One main difference between the SA and PZ models is that in the SA model neighboring β sheets are tightly packed at 8.3 Å (a distance that is crucial to explain the diffraction pattern of $D_2Q_{15}K_2$), while in the PZ model, the intersheet distance is 16.8 Å, which leaves a considerable gap between the side chains of the stacking sheets. More recent evidence confirming the existence of the three possible types of hydrogen bonds was provided by NMR studies[34] and MD simulations,[46] although these studies did not report on the explicit conformation of the side chains.

Our simulations are consistent with the bulk of this data but also provide further insight into the mechanisms related to the side-chain packing. In agreement with the model proposed by SA, the β sheets are tightly interdigitated. However, in order to arrive at this configuration, the authors were forced to flip successive side-chain N$-C_\alpha-C_\beta-C_\gamma$ torsion angles by 180°, because they estimated that the PZ mechanism would fail to establish a continuous hydrogen-bond pattern (Figure 5 in ref 32). In addition, in order to achieve agreement with the diffraction pattern, they had to displace the intermediate layers by 0.6*a*. (Although not explicitly stated, from the look of their figures it appears as if they used the TD APS-APP arrangement for the antiparallel sheets and the ZD PS, either PP or APP, arrangement for the parallel sheets). Our simulations show that

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

the N−C_α−C_β−C_γ flipping is neither favored nor necessary for the *continuous* hydrogen bonding of the side chains, and this can simply be achieved with the much cheaper rotation of the terminal groups in the side chains (i.e., the C_β−C_γ−C_δ−N_ε angle). In our simulations, the system is allowed to relax on its own, and several initial conditions converge in a quarter-stagger conformation (0.25*a*), which is in agreement with experimental observations.[41] This results in a cell of smaller area than the original one, with a C_α−C_α distance of 8.25 Å for the APS systems. As shown in Table 2, the final lattice parameters are consistent with crystal parameters that can be inferred from X-ray fiber diffraction patterns. One could still expect small differences, since in general the experiments use slightly different peptides, such as $D_2Q_{15}K_2$,[28] which has four charged groups at the end and probably a turn conformation. These diffraction patterns however lack fully atomic resolution and therefore would probably not be able to discern between the different lattice structures in this work. As pointed out by SA, both PS and APS would give rise to virtually the same diffraction pattern (except for technically altering the Miller indices, which were not provided in the original diffraction paper), but the PS arrangement is "less satisfactory". This seems to be the case for our simulations, where the lattice parameters (Table 2 and Supporting Information Figure S5) are slightly closer to the diffraction values in the APS than the PS systems. In addition, considerations of the stability of smaller systems is also important. As discussed in the previous sections, smaller aggregates are not stable for PS systems, and ZD systems are in general more stable that those that exhibit TD. Thus, one would expect that the initial stability of, say, dimers might exert a bias toward higher order aggregates of the same nature.

Structure and Stability of the Q_{40} Monomeric Forms. The main results can be summarized as follows: The most compact models, with four β strands of approximately seven residues each and interdigitated side chains, are stable in the 1 μs time scale. This makes them excellent candidates for the monomeric misfolded structure capable of triggering aggregation. By contrast, the single-sheet 4 × 1 β structure connected by three turns is unstable, despite having the largest initial number of residues in a β conformation. In particular, the lateral β sheet stack (model 1), the β double hook (model 2), and the collapsed β helix (model 4) are more stable than the diagonal β sheet stack (model 3). These results confirm the previous conclusion that compact structures are more stable than less compact ones. Out of the three stable models, the lateral β sheet stack (model 1) seems to be slightly "overly stable", because it is still stable with 33 residues, while the collapsed β helix (model 4) undergoes a sharp transition in its stability somewhere between 37 and 33 residues. Thus, it would appear that model 4 best describes the onset of a threshold around 35−36 residues.

A previous simulation[46] for a Q_{57} collapsed β helix with interdigitated side chains was the first one to confirm the stability of this type of aggregate (up to 50 ns), but the authors noted having difficulties in the equilibration of the equivalent Q_{40} aggregate. Indeed, their coordinates were used for another stability study carried out by a different group that found this type of aggregate unstable.[48] This second work found that the most stable structure, in addition to an α helix, was the single 4 × 1 β sheet.[48] Our conclusions are in stark contrast to this second work. First, their conclusions were based on 100 ns simulations. We find that the single β sheet is the most unstable structure, but the instability effects only start showing up after about 500 ns. Second, our study brings out the importance of the side-chain conformations for stable protofibrils or crystallites. Although the authors in this previous work do not describe the conformation of the side chains, a careful look at their initial conditions would suggest that these do not completely satisfy preferred side-chain conformations. This is not surprising, because the initial side-chain coordinates were generated[46] using the coordinates of crystal structures of GNNQQNY, which does not necessarily result in ideal polyQ packing. In this sense, our results revindicate the conclusions of ref 46, which proposed this structure as a good candidate for the nucleated monomer, at least for Q_{57}. We show that with the appropriate side-chain conformations, a Q_{40} monomer can indeed be stable, and this would therefore agree with the threshold limit for aggregation.

There have been several experiments that try to measure the relative stability of arcs versus turns,[35,36] but these studies generally consider shorter polyQ stretches, generally 23−24 residues, with or without turn-enhancing amino acid inclusions in the middle, which result in either only one arc or only one turn. These results favor the stability of the single β hairpin, but it is not clear how these studies would translate to the case of Q_{40}, where a total of three arcs or hairpins are present.

METHODS

Oligomeric structures of Ace-$(Gln)_8$-Nme. Here we describe how the initial oligomers of Ace-$(Gln)_8$-Nme (Q_8) were constructed. First, a Q_8 peptide was generated by using the AMBER 12 package[49] with the ff12SB force field. In this first instance, the peptide is created with all its torsion angles equal to 180°. In order to create parallel and antiparallel strands, the backbone dihedral angles of consecutive monomers were rotated by using the PyMOL package[50] to values (−120°,115°) for parallel and (−140°,135°) for antiparallel conformations. Then the peptide was rotated so that its principal axis became parallel to the *z* axis and the C_α atoms lay on the *yz* plane. The side chains were then rotated so that that they lay parallel to the *y* axis. This affected the torsion angle N−C_α−C_β−C_γ, while leaving the other side-chain torsion angles unaffected and equal to their 180° initial value. Different initial values of the N−C_α−C_β−C_γ angle were used as initial conditions, as discussed in the Results section. Two Q_8 monomers were then brought together so that the peptide axes were along the *z*-axis, the hydrogen bonds along the *x*-axis (so that the β sheet was on the *x*−*z* plane), and the side chains were perpendicular to the β-sheet plane, along the *y*-axis. The distance between the C_α atoms of *i*th and (*i* + 2)th positions in a peptide was c = 0.657 nm for parallel and c = 0.695 nm for antiparallel conformations. We then proceeded to construct the $M \times N$ oligomeric "crystals", as discussed in the Results section.

Monomeric Q_{40} Protofibril. Given that the observed β-strand length in a fibril is seven to eight residues,[14,31,39] a Q_{40} peptide can give rise to a maximum of four strands, including the connecting turns and arcs. Based on the results obtained for the oligomeric structures of Ace-$(Gln)_8$-Nme, in order to build the monomeric Q_{40} protofibril structures, we first took 2 × 2 subsystems, containing four central Q_8 peptides, that belong to the relaxed and very stable structures of the 4 × 4 ZD systems at 1 μs in the quarter-stagger[41] conformations. The four short peptides were connected into a long peptide by arc turns or hairpin turns or a combination of arc and hairpin turns. A hairpin peptide that joins two Q_8 peptides in the same sheet was constructed by adding a Q residue at the turn position, while an arc turn that joins two peptides in two different sheets was constructed by adding a Q_3 peptide at the arc position.

Given that the addition of arcs and hairpins changes the number of total residues, before adding a turn to a Q_8 strand, we added one residue (such that it became a Q_9 strand) or occasionally removed one residue (therefore obtaining a Q_7 strand), such that all aggregates have a total of 40 residues. All the Q_{40} models shown in Figure 9 are capped by Ace- at the amino end and by -Nme at the carboxyl end. A Q_8 strand can be expanded to a Q_9 strand by adding an N- or a C-terminal residue. For a C-expansion, we consider a vector that connects the C atoms of the 6th and 8th residues and use this vector to translate a copy of the 7th residue

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

and add it to the C-terminus of the Q_8 strand. A similar definition (using the 3rd and 1st residues to define the vector) is used to add a copy of the 2nd residue to the N-terminus of the strand. The hairpin (Figure S13, Supporting Information) is then built as follows. A suitable glutamine residue in coil or turn structure is picked from a 100 ns MD simulation of the monomeric Q_{40} system. The two atoms to be connected are the C atom at the C-terminus of the first peptide (C^1) and the N atom at the N-terminus of the second peptide (N^2), whose distance is 4.1−4.6 Å. Given that the peptide bond distance is in the range 1.32−1.33 Å, the distance between the N and C atoms of the added residue should be larger than 2.0 Å. The residue is then translated and rotated so that it fits in the available space without any bond stretching and the side chain is on the plane pointing outward, away from the steric zipper. The creation of the arc turn (Figure S14, Supporting Information) is similar to that of the hairpin case. The Q_3 fragment was chosen from a 100 ns MD simulation of the monomeric Q_{40} system in such a way that it properly fits the ∼8.3 Å distance between the two peptides to be joined. The selected curved fragment has the three side chains on the convex side and is conveniently translated and rotated to fit the gap, as shown in Figure S14, Supporting Information.

Simulation Details. The simulations were carried out using the AMBER 12 package[49] with the ff12SB force field and the explicit TIP3P water model. Each aggregate was placed in a cubic box with a side of approximately 6 nm under periodic boundary conditions, with over 6680 waters. The distance from the peptides to the box edge was kept at a minimum of 1.0 nm to prevent the peptides from interacting with their own images. An 8 Å cutoff was applied for the nonbonded interactions, and the particle mesh Ewald (PME) method[51] was used to calculate the electrostatic interactions with cubic-spline interpolation and a grid spacing of approximately 1 Å. The equilibration of the solvated systems followed five steps. First, a short steepest descent minimization was executed by applying the conjugate gradient method and restraining the peptide atoms to their initial positions. Second, an unrestrained steepest descent method followed by conjugate gradient minimization was performed. Third, the system was heated from 0 to 310 K under constant volume conditions and weak restraints on the peptide atoms. Fourth, 100 ps NPT runs with weak restraints on the peptide atoms were carried out to reach stable density (around 1.0 g/cm^3). Fifth, 100 ps NPT Langevin dynamics with constant pressure at 310 K without restraints were carried out.

The final configurations for the Q_8 oligomers were used for 100 ns production runs for all aggregates except the 4 × 4 ones, which were run until 1 μs (ZD) and 1.3−1.5 μs (TD). The five Q_{40} configurations were run up to 1 μs each. In addition, several Q_n monomer simulations with n near the 35 threshold were simulated for various monomer models for a total time of 2.7 μs. This results in a cumulative simulation time of over 23 μs. Configurations were saved every 100 ps.

■ ASSOCIATED CONTENT

Supporting Information

Populations and broken time of the P_{bs}^{2s}, P_{bs}^{3s}, and P_{bs}^{4s} states of the single sheet models, population of states as a function of the number of extended residues and time dependence of the β content of the ZD 3 × 1 Q_8 systems, population of states as a function of the number of extended residues of the ZD 3 × 2, ZD 2 × 3, ZD 4 × 2, and ZD 2 × 4, possible hydrogen bonds for the 4 × 1, 2 × 2, 2 × 3, 3 × 2, 2 × 4, and 4 × 2 systems, time dependence of the crystal parameters in 4 × 4 systems, time dependence and distribution of the torsion angles of the central residues in 4 × 4 systems, and illustrations of the connection of two strands by a hairpin turn or by an arc turn. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Celeste Sagui. E-mail: sagui@ncsu.edu.

Author Contributions

V.H.M. ran simulations; V.H.M. and C.S. analyzed data; all authors helped formulate the problem and wrote paper.

Funding

This work has been supported by National Science Foundation (NSF) via Grant SI2-SSE-1148144 and MCB grants.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Robert Rose for useful discussions and the NC State HPC Center for extensive computational support.

■ REFERENCES

(1) Zoghbi, H. Y., and Orr, H. T. (2000) Glutamine Repeats and Neurodegeneration. *Annu. Rev. Neurosci. 23*, 217−247.

(2) Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003) Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis. *Science 300*, 486−489.

(3) Kar, K., Jayaraman, M., Sahoo, B., Kodali, R., and Wetzel, R. (2011) Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nat. Struct. Mol. Biol. 18*, 328−336.

(4) Scherzinger, E., Sittler, A., Schweiger, K., Heiser, V., Lurz, R., Hasenbank, R., Bates, G. P., Lehrach, H., and Wanker, E. E. (1999) Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: Implications for Huntington's disease pathology. *Proc. Natl. Acad. Sci. U. S. A. 96*, 4604−4609.

(5) Chen, S., Berthelier, V., Yang, W., and Wetzel, R. (2001) Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. *J. Mol. Biol. 311*, 173−182.

(6) Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S. W., and Bates, G. P. (1996) Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell 87*, 493−506.

(7) DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., and Aronin, N. (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science 277*, 1990−1993.

(8) Poirier, M. A., Li, H., Macosko, J., Cai, S., Amzel, M., and Ross, C. A. (2002) Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. *J. Biol. Chem. 277*, 41032−41037.

(9) Wacker, J. L., Zareie, M. H., Fong, H., Sarikaya, M., and Muchowski, P. J. (2004) Hsp70 and Hsp40 attenuate formation of spherical and annular polyglutamine oligomers by partitioning monomer. *Nat. Struct. Mol. Biol. 11*, 1215−1222.

(10) Lee, C. C., Walters, R. H., and Murphy, R. M. (2007) Reconsidering the mechanism of polyglutamine peptide aggregation. *Biochemistry 46*, 12810−12820.

(11) Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) Huntingtin-Encoded Polyglutamine Expansions Form Amyloid-like Protein Aggregates In Vitro and In Vivo. *Cell 90*, 549−558.

(12) Chen, S., Ferrone, F. A., and Wetzel, R. (2002) Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc. Natl. Acad. Sci. U. S. A. 99*, 11884−11889.

(13) Moradi, M., Babin, V., Roland, C., and Sagui, C. (2012) Are long-range structural correlations behind the aggregation phenomena of polyglutamine diseases? *PLoS Comput. Biol. 8*, No. e1002501.

(14) Thakur, A. K., and Wetzel, R. (2002) Mutational analysis of the structural organization of polyglutamine aggregates. *Proc. Natl. Acad. Sci. U. S. A. 99*, 17014−17019.

(15) Muchowski, P. J., and Wacker, J. L. (2005) Modulation of neurodegeneration by molecular chaperones. *Nat. Rev. Neurosci. 6*, 11−22.

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645

(16) Nagai, Y., Inui, T., Popiel, H. A., Fujikake, N., Hasegawa, K., Urade, Y., Goto, Y., Naiki, H., and Toda, T. (2007) A toxic monomeric conformer of the polyglutamine protein. *Nat. Struct. Mol. Biol. 14*, 332−340.
(17) Miller, J., Arrasate, M., Shaby, B. A., Mitra, S., Masliah, E., and Finkbeiner, S. (2010) Quantitative Relationships between Huntingtin Levels, Polyglutamine Length, Inclusion Body Formation, and Neuronal Death Provide Novel Insight into Huntington's Disease Molecular Pathogenesis. *J. Neurosci. 30*, 10541−10550.
(18) Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature 431*, 805−810.
(19) Kopito, R. (2000) Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol. 10*, 524−530.
(20) Ross, C. A., and Poirier, M. A. (2005) What is the role of protein aggregation in neurodegeneration? *Nat. Rev. Mol. Cell Biol. 6*, 891−898.
(21) Bodner, R. A., Outeiro, T. F., Altmann, S., Maxwell, M. M., Cho, S. H., Hyman, B. T., McLean, P. J., Young, A. B., Housman, D. E., and Kazantsev, A. G. (2006) Pharmacological promotion of inclusion formation: A therapeutic approach for Huntington's and Parkinson's diseases. *Proc. Natl. Acad. Sci. U. S. A. 103*, 4246−4251.
(22) Waelter, S., Boeddrich, A., Lurz, R., Scherzinger, E., Lueder, G., Lehrach, H., and Wanker, E. E. (2001) Accumulation of mutant huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation. *Mol. Biol. Cell 12*, 1393−1407.
(23) Taylor, J. P., Tanaka, F., Robitschek, J., Sandoval, C. M., Taye, A., Markovic-Plese, S., and Fischbeck, K. H. (2003) Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein. *Hum. Mol. Genet. 12*, 749−757.
(24) Glabe, C. G., and Kayed, R. (2006) Common structure and toxic function of amyloid oligomers implies a common mechanism of pathogenesis. *Neurology 66*, S74−S78.
(25) Armen, R. S., Bernard, B. M., Day, R., Alonso, D. O. V., and Daggett, V. (2005) Characterization of a possible amyloidogenic precursor in glutamine-repeat neurodegenerative diseases. *Proc. Natl. Acad. Sci. U. S. A. 102*, 13433−13438.
(26) Tomita, K., Popiel, H. A., Nagai, Y., Toda, T., Yoshimitsu, Y., Ohno, H., Oishi, S., and Fujii, N. (2011) Structure-activity relationship study on polyglutamine binding peptide QBP1. *Bioorg. Med. Chem. 17*, 1259−1263.
(27) Fiumara, F., Fioriti, L., Kandel, E. R., and Hendrickson, W. A. (2010) Essential role of coiled coils for aggregation and activity of Q/N-rich prions and PolyQ proteins. *Cell 143*, 1121−1135.
(28) Perutz, M. F., Finch, J. T., Berriman, J., and Lesk, A. (2002) Amyloid fibers are water-filled nanotubes. *Proc. Natl. Acad. Sci. U. S. A. 99*, 5591−5595.
(29) Daggett, V. (2006) α-Sheet: The Toxic Conformer in Amyloid Diseases? *Acc. Chem. Res. 39*, 594−602.
(30) Babin, V., Roland, C., and Sagui, C. (2011) The α-sheet: A missing-in-action secondary structure? *Proteins: Struct., Funct., Bioinf. 79*, 937−946.
(31) Sharma, D., Shinchuk, L. M., Inouye, H., Wetzel, R., and Kirschner, D. A. (2005) Polyglutamine homopolymers having 8−45 residues form slablike beta-crystallite assemblies. *Proteins 61*, 398−411.
(32) Sikorski, P., and Atkins, E. (2005) New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils. *Biomacromolecules 6*, 425−432.
(33) Makin, O. S., Atkins, E., Sikorski, P., Johansson, J., and Serpell, L. C. (2005) Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. U. S. A. 102*, 315−320.
(34) Schneider, R., Schumacher, M. C., Mueller, H., Nand, D., Klaukien, V., Heise, H., Riedel, D., Wolf, G., Behrmann, E., Raunser, S., Seidel, R., Engelhard, M., and Baldus, M. (2011) Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. *J. Mol. Biol. 412*, 121−136.
(35) Buchanan, L. E., Carr, J. K., Fluitt, A. M., Hoganson, A. J., Moran, S. D., de Pablo, J. J., Skinner, J. L., and Zanni, M. T. (2014) Structural motif of polyglutamine amyloid fibrils discerned with mixed-isotope infrared spectroscopy. *Proc. Natl. Acad. Sci. U. S. A. 111*, 5796−801.
(36) Kar, K., Hoop, C. L., Drombosky, K. W., Baker, M. A., Kodali, R., Arduini, I., van der Wel, P. C., Horne, W. S., and Wetzel, R. (2013) β-hairpin-mediated nucleation of polyglutamine amyloid formation. *J. Mol. Biol. 425*, 1183−1197.
(37) Nelson, R., Sawaya, M. R., Balbirnie, M., ϕMadsen, A., Riekel, C., Grothe1, R., and Eisenberg, D. (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature 435*, 773−778.
(38) Buchete, N.-V., Tycko, R., and Hummer, G. (2005) Molecular Dynamics Simulations of Alzheimer's β-Amyloid Protofilaments. *J. Mol. Biol. 353*, 804−821.
(39) Darnell, G., Orgel, J. P., Pahl, R., and Meredith, S. C. (2007) Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure. *J. Mol. Biol. 374*, 688−704.
(40) Krull, L. H., Wall, J. S., Zobel, H., and Dimler, R. J. (1965) Synthetic Polypeptides Containing Side-Chain Amide Groups: Water-insoluble Polymers. *Biochemistry 4*, 626−633.
(41) Fraser, R. D. B., and MacRae, T. P. (1973) *Conformation in Fibrous Proteins and Related Synthetic Polypeptides*, Academic Press, New York.
(42) Eanes, E. D., and Glenner, G. G. (1968) X-ray diffraction studies on amyloid filaments. *J. Histochem. Cytochem. 16*, 673−677.
(43) Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol. 273*, 729−739.
(44) Tycko, R. (2006) Molecular structure of amyloid fibrils: insights from solid-state NMR. *Q. Rev. Biophys. 39*, 1−55.
(45) Nakano, M., Ebina, K., and Tanaka, S. (2013) Study of the aggregation mechanism of polyglutamine peptides using replica exchange molecular dynamics simulations. *J. Mol. Model. 19*, 1627−1639.
(46) Esposito, L., Paladino, A., Pedone, C., and Vitagliano, L. (2008) Insights into Structure, Stability, and Toxicity of Monomeric and Aggregated Polyglutamine Models from Molecular Dynamics Simulations. *Biophys. J. 94*, 4031−4040.
(47) Perutz, M. F., Johnson, T., Suzuki, M., and Finch, J. T. (1994) Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc. Natl. Acad. Sci. U. S. A. 91*, 5355−5358.
(48) Miettinen, M. S., Knecht, V., Monticelli, L., and Ignatova, Z. (2012) Assessing Polyglutamine Conformation in the Nucleating Event by Molecular Dynamics Simulations. *J. Phys. Chem. B 116*, 10259−10265.
(49) Case, D. A., et al. (2012) *AMBER 12*; University of California, San Francisco.
(50) Schrödinger, L. L. C., http://www.pymol.org.
(51) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577.

DOI: 10.1021/cn500358g
ACS Chem. Neurosci. 2015, 6, 632−645