Transformation between $\alpha$-helix and $\beta$-sheet structures of one and two polyglutamine peptides in explicit water molecules by replica-exchange molecular dynamics simulations

Hsin-Lin Chiang,* Chun-Jung Chen†, Hisashi Okumura‡, Chin-Kun Hu§

April 19, 2014

Abstract

Aggregation of polyglutamine peptides with $\beta$-sheet structures is related to some important neurodegenerative diseases such as Huntington’s disease. However it is not clear how polyglutamine peptides form the $\beta$-sheets and aggregate. To understand this problem, we performed all-atom replica-exchange molecular dynamics (REMD) simulations of one and two polyglutamine peptides with ten glutamine residues in explicit water molecules. Our results show that two polyglutamine peptides mainly formed helix or coil structures when they are separated, as in the system with one-polyglutamine peptide. As the inter-peptide distance decreases, the intra-peptide $\beta$-sheet structure sometimes appear as an intermediate state, and finally the inter-peptide $\beta$-sheets are formed. We also find that the polyglutamine dimer tends to form the anti-parallel $\beta$-sheet conformations rather than the parallel $\beta$-sheet, which is consistent with previous experiments and a coarse-grained molecular dynamics simulation.

Keywords: molecular dynamics simulation, neurodegenerative diseases, protein aggregation, polyglutamine

*Department of Physics, National Tsing Hua University - Hsinchu 30013, Taiwan. Institute of Physics, Academia Sinica - Taipei 11529, Taiwan.
†Department of Physics, National Tsing Hua University - Hsinchu 30013, Taiwan. Life Science Group, Scientific Research Division, National Synchrotron Radiation Research Center - Hsinchu 30076, Taiwan. Institute of Biotechnology and University Center for Bioscience and Biotechnology, National Cheng Kung University - Tainan City 701, Taiwan.
‡Research Center for Computational Science, Institute for Molecular Science - Okazaki, Aichi 444-8585, Japan. Department of Structural Molecular Science, The Graduate University for Advanced Studies - Okazaki, Aichi 444-8585, Japan. E-mail: hokumura@ims.ac.jp
§Institute of Physics, Academia Sinica - Taipei 11529, Taiwan. E-mail: huck@mails.phys.sinica.edu.tw
Transformation pathways and peptide conformations of a system with two polyglutamine peptides and 3129 water molecules obtained by replica-exchange molecular dynamics simulations. When the inter-peptide distance gets shorter, α-helix structure decreases and the inter-peptide β-sheet structure increases. The intra-peptide β-sheet structure appears as an intermediate state in the α-helix to inter-peptide β-sheet transformation.
INTRODUCTION

Neurodegenerative diseases such as Huntington’s disease (HD), Alzheimer’s disease, spinocerebellar ataxia, and type II diabetes are related to protein aggregation. Such diseases become more serious in longevity societies. Thus the study of protein or peptide aggregation has attracted a lot of attention from researchers in the medical, biochemical, and biophysical communities. The symptoms of these diseases might be different and they might appear in different organisms. However, different precursor proteins of many diseases may share similar morphological features in aggregation or amino acid sequence. For example, Huntington’s disease, spinobulbar muscular atrophy, dentatorubral-pallidoluysian atrophy, and six kinds of spinocerebellar ataxias are well known as trinucleotide repeat disorder of inherited neurodegenerations. Theses diseases are considered to be caused by abnormal polyglutamine (polyQ) repeat encoded by expanded CAG repeats in genomic sequence. Huntington’s disease is the most common among these diseases which has prevalence estimated about 4 in 10,000 people of western European ancestry.

It is well known that people with polyQ stretches of 38 residues or longer carry high risk of HD. In 1998, Myers, Marans and MacDonald reported that age-of-onset correlates with repeat length: repeat lengths of 38-39 are associated with onset in the 60-80 age range, while repeat lengths longer than 70 lead to disease with onset in the 5-15 age range. In 2002 Chen et al. studied the in vitro aggregation kinetics of a series of polyQ peptides. They used the kinetic parameters to predict aggregation curves for very low concentrations of polyQ that might appear in the cell. They found that the dependence of the differences in aggregation lag times on repeat-length is similar to the dependence of the differences in age-of-onset on repeat-length in HD patients. Although many researches indicate 36-40 repeated glutamines trigger amyloid fibril formation in vivo, polyQs as short as 6-8 residues have also been found to aggregate in experiments.

In 2002 Perutz et al. proposed a water-filled nanotube model (WFNM) to explain x-ray diffraction pattern from the deposit of aggregated polyglutamine (D$_2$Q$_{15}$K$_2$). In this model, molecules of D$_2$Q$_{15}$K$_2$ form helical structures with circular cross section and neighboring molecules lie in parallel to each other. This model has also been proposed as the common
molecular structure of amyloidoses by Singer and Dewji’s in 2006. However, Stork et al. found from their simulations that the β-sheet with triangular cross section is more stable than that with circle cross section.

In 2005 Sikorski and Atkins (SA) used an alternative model to reproduce Perutz’s diffraction pattern. In SA model, polyQ peptides assemble to form crystal as a typical cross-β structure, in which polyQ peptides fold as β-hairpins and then stack to form cross-structure preferring anti-parallel β-sheets in terms of hydrogen-bonding interactions. Thus in WFN, neighboring polyQ peptides are parallel, while in SA model, neighboring polyQ peptides are anti-parallel. In 2007, Darnell et al. used experimental method to find that flanking residues affect dimerization and aggregation.

In 2010 Laghaei and Mousseau used replica-exchange molecular dynamics (REMD) method and optimized potential for effective peptide (OPEP) force field to study a dimer of polyQ peptides with 40 residues and get an anti-parallel cylindrical structure. It should be mentioned that some references included Perutz’s earlier paper had also suggested that polyQ peptides tend to form anti-parallel β-sheet conformations. In 2008 Ogawa et al. used all-atom force field with explicit water and proved the stability of polyQ using a β-helix as the initial structure. In 2011 Dlugosz and Trylska performed all-atom REMD in implicit solvent and MD in explicit solvent to prove that polyQ tracts connected with Htt N-terminal fragment mainly adopt α-helix conformations. In 2013 Nakano et al. performed REMD simulations in implicit water and showed polyQ prefer anti-parallel β-sheet.

All-atom force field in explicit water is more reliable than in implicit water although it costs expensive computer resources. To clarify whether parallel or anti-parallel β-sheet is more stable, here we perform REMD simulations of systems with one and two polyQ peptides in 1611 and 3129 explicit water molecules, respectively. For a given simulation time interval, REMD simulations can sample more protein conformations in wider parameter space and thus can give more accurate free energy landscape than conventional MD simulations. Such accurate free energy landscapes are useful for deducing folding or aggregation pathways of proteins although REMD cannot give dynamic behavior of the system in real time. Our study is the first REMD simulations in which the aggregation of polyQ peptides is studied in all-atom force field with explicit water.
We find that the polyglutamine dimer tends to form the anti-parallel $\beta$-sheet conformations rather than the parallel $\beta$-sheet, which is consistent with previous experiments$^{24,33}$ and a coarse-grained molecular dynamics simulation$^9$.

METHODS

We performed REMD simulations of one and two polyQ-peptide systems. Each polyQ peptide consists of 10 repeated glutamines. The N terminus and C terminus of the polyQ peptides were capped by an acetyl group and N-methyl group, respectively. In the one polyQ-peptide system, we put one polyQ peptide in a cubic simulation box with the length $L = 37.729$ Å and filled the box with 1611 water molecules. In the two polyQ-peptide system, we put two polyQ peptides in a cubic simulation box with the length $L = 46.608$ Å and filled the box with 3129 water molecules. The initial condition was prepared without any secondary structure. The backbone dihedral angles $\phi$ and $\psi$ were started from $\phi = \psi = 180^\circ$. The polyQ peptides were placed so that the angle between the two peptides will be perpendicular. The shortest distance among peptides was set to 10 Å.

We employed the GEMB (Generalized-Ensemble Molecular Biophysics) program developed by Okumura$^{45–49}$ to perform REMD simulations. We used Amber 99SB$^{50}$ force field for the peptides and TIP3P$^{51}$ rigid-body model for the water molecules. The electrostatic potential was calculated using the Particle Mesh Ewald (PME) method. The cutoff distance was $r_c = 12$ Å for the Lennard-Jones potential. The symplectic rigid body molecular dynamics algorithm$^{52}$ was used for the water molecules. Reversible multiple time scale molecular dynamics$^{53}$ techniques were also applied. The time step was $\Delta t = 0.5$ fs for the protein atoms and $\Delta t = 4.0$ fs for the water molecules. Because the symplectic rigid-body algorithm was used for the water molecules here, $\Delta t$ taken was as long as 4.0 fs$^{45}$. The Nosé-Hoover thermostat was used to obtain canonical ensemble$^{54–56}$.

The equations of motion of the system are given by$^{45,47,52}$

\begin{align}
\dot{r}_i &= \frac{p_i}{m_i}, \\
\dot{p}_i &= F_i - \zeta p_i,
\end{align}

5
\[ q_i = \frac{1}{2} \vec{S}(q_i) \omega_i^{(4)}, \]  
\[ \vec{I}_i \dot{\omega}_i = N_i - \omega_i \times (\vec{I}_i \omega_i) - \zeta \vec{I}_i \omega_i, \]  
\[ \dot{\zeta} = \frac{1}{Q} \left[ \sum_{i=1}^{N+M} \frac{p_i}{m_i} + \sum_{i=1}^{M} \omega_i^T \vec{I}_i \omega_i - gk_B T_0 \right], \]

where index \( i \) is taken for both \( N \) peptide atoms and \( M \) rigid-body water molecules. When \( i \) indicates a peptide atom, \( r_i \) stands for its coordinate; when \( i \) indicates a rigid-body water molecule, \( r_i \) stands for the coordinate of its center of mass (COM). The vector \( p_i \) and the constant \( m_i \) are the momentum and the mass, respectively, of atom \( i \) or rigid body molecule \( i \).

The constant \( Q \) is the artificial “mass” of the Nosé-Hoover thermostat. The vector \( F_i \) stands for the force acting on atom \( i \) if \( i \) indicates an atom in the peptide molecules or the total force on all the atoms in molecule \( i \) if \( i \) indicates a rigid-body water molecule. The variable \( \zeta \) is the momentum-dependent friction coefficient of the Nosé-Hoover thermostat. The variable \( q = (q_0, q_1, q_2, q_3)^T \) is a quaternion which specifies the orientation of the rigid-body molecule. The elements of the matrix \( \vec{S}(q) \) are given by:

\[ \vec{S}(q) = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix}. \]  

The vector \( \omega \) is the angular velocity, whose elements are given by

\[ \omega = (\omega_1, \omega_2, \omega_3)^T , \]  
and \( \omega^{(4)} \) is the four-dimensional angular velocity:

\[ \omega^{(4)} = (0, \omega_1, \omega_2, \omega_3)^T . \]  

The tensor \( \vec{I} \) is the principal inertia tensor. The vector \( N_i \) is the torque of molecule \( i \), which is calculated by

\[ N_i = \sum_{\alpha \in i} r_\alpha \times F_\alpha , \]  
where the summation is over all atoms (labeled by \( \alpha \)) of molecule \( i \), \( r_\alpha \) and \( F_\alpha \) are the coordinate of atom \( \alpha \) and force acting on atom \( \alpha \), respectively. The constant \( T_0 \) is the set
temperature for the Nosé-Hoover thermostat and $g$ is the number of degrees of freedom: $g = 3N + 6M$.

In REMD simulation, the temperatures are exchanged between replicas with neighboring temperatures from $(x_{m}^{i}, x_{n}^{i})$ to $(x_{n}^{i}, x_{m}^{i})$ depending on the Metropolis criterion,

\[
w(x_{m}^{i}, x_{n}^{i} | x_{m}^{j}, x_{n}^{j}) = \begin{cases} 
1 & \text{if } \Delta \leq 0, \\
\exp(-\Delta) & \text{if } \Delta > 0.
\end{cases}
\]

where $x_{m}^{i}$ and $x_{n}^{j}$ are coordinate sets of replicas $i$ and $j$ with temperature $T_{m}$ and $T_{n}$, respectively. The difference $\Delta$ is given by $\Delta = (1/k_{B}T_{n} - 1/k_{B}T_{m})(E_{i} - E_{j})$ and $k_{B}$ is the Boltzmann constant. $E_{i}$ and $E_{j}$ are potential energies of $i$ and $j$ replicas, respectively. We tried to exchange the temperatures every 1 ps.

In the one polyQ-peptide system, we used 17 temperatures at 300.0, 303.8, 307.7, 312.0, 316.5, 321.3, 326.5, 331.9, 337.8, 343.9, 350.5, 357.5, 365.0, 373.0, 381.4, 390.4, 400.0 K. In the two polyQ-peptide system, we used 24 temperatures for REMD at 300, 302.6, 305.3, 308.1, 311.1, 314.1, 317.3, 320.7, 324.2, 327.9, 331.7, 335.7, 339.9, 344.2, 348.8, 353.5, 358.5, 363.7, 369.1, 374.7, 380.7, 386.8, 393.3, and 400.0 K. Note that the highest and lowest temperatures are the same in both system. Because the system size of the one polyQ-peptide system is smaller than the two peptide system, fewer number of replicas was enough to exchange the temperatures. The average acceptance ratio was 36% in the one peptide simulation. It was 27% in the two peptide system.

We first performed REMD simulations for 10 ns per replica as equilibration processes. We then performed 100 ns simulations per replica for data collection to calculate physical quantities. The total simulation time for the conformation sampling was 1.7 $\mu$s for the one polyQ-peptide system and 2.4 $\mu$s for the two polyQ-peptide system.

The definition of the secondary structure of proteins (DSSP)\textsuperscript{57} was used to identify the secondary structures of polyQ peptides. According to DSSP, hydrogen bonds are defined with the equation

\[
E = 332q_{1}q_{2}\left(\frac{1}{r_{ON}} + \frac{1}{r_{CH}} - \frac{1}{r_{OH}} - \frac{1}{r_{CN}}\right)\text{kcal/mol},
\]

where $q_{1} = 0.42$ and $q_{2} = 0.2$. The variables $r_{ON}$, $r_{CH}$, $r_{OH}$, and $r_{CN}$ are distances between two atoms in two amino acid residues in the unit Å. Subscripts O, N, C, and H mean...
oxygen, nitrogen, carbon, and hydrogen atoms, respectively. A hydrogen bond is identified if the potential energy $E$ calculated from eq. (11) is smaller than -0.5 kcal/mol. By defining hydrogen bonds, bridge and turn can be decided further, which are followed by the definitions of α-helix, 3$_{10}$-helix, π-helix, and β-sheet.

To compare the number of secondary structures fairly between the helix and β-sheet structures, the priority of secondary structure in DSSP was adjusted in this paper from H>B>E>G>I to H>G>B, here H is a abbreviation of α-helix, G is 3$_{10}$-helix, I is π-helix, B is β-ladder, and E is extended β-ladder (B and E are both considered to be β-sheet). This adjustment means that if one residue was considered to be α-helix, β-sheet, and 3$_{10}$ helix simultaneously, we labeled it not only α-helix but also β-sheet. This is different from the official DSSP program, which labeled it only α-helix because of its highest priority.

RESULTS and DISCUSSION

The free energy landscape of one polyQ-peptide system at $T = 300$ K is shown in Fig. 1. The definition of the free energy difference $\Delta F$ as a function of reaction coordinates ($\xi_1, \xi_2$) is given by

$$\Delta F(\xi_1, \xi_2) = -kT \ln P(\xi_1, \xi_2) - F_{\text{min}}$$

where $P(\xi_1, \xi_2)$ is the probability distribution of $\xi_1$ and $\xi_2$. For the purpose to investigate the transformation between helix and β-sheet structures and to distinguishing the parallel and anti-parallel β-sheet, we chose the length of helix and number of β-bridges as the reaction coordinates $\xi_1$ and $\xi_2$ in Fig. 1. Here, helix structures include α-helix, 3$_{10}$-helix, and π-helix structures. $F_{\text{min}}$ is the minimum value of $-kT \ln P(\xi_1, \xi_2)$.

We chose some representative states and labeled them from states A to I as shown in Fig. 1. The snapshots of states A to I were created by RasMol and shown in Fig. 2. The relative positions of states A to I in Fig. 2 are the same as those in Fig. 1. REMD simulation can overcome the free-energy barriers and one state can move to neighboring states. Here, the neighboring state is a state that is close from the current state. Some transformation pathways between the neighboring states are also shown using arrows in Fig. 2. Note that not only these pathways but also other pathways may be possible. One state can move to a
neighboring state through another state. The global minimum state is B, which is a random coil state. State B can be reached directly from neighboring states A, C, and F. State A is a parallel intra-peptide $\beta$-sheet. States C, D, and E have one, two, and three anti-parallel intra-peptide $\beta$-sheets (i.e. $\beta$-hairpin), respectively. According to the DSSP, the minimum length of the helix is three of the $3_{10}$-helix structure, which is observed at state F. With an additional turn and displacement of the helix residues, state F transforms to state G with a $3_{10}$-helix and one anti-parallel $\beta$-bridge. State H has five helix residues, which include $3_{10}$-helix, $\alpha$-helix, and $\pi$-helix structures. The length of helix can extend up to eight, as illustrated at state I. This is the longest helix structure for a ten-residue peptide because the N-terminus and C-terminus residues cannot be determined as a helix structure by the DSSP.

We can also see in Fig. 1 and Fig. 2 that the anti-parallel $\beta$-sheet structure has higher proportion than the parallel $\beta$-sheet structure. The probabilities of the helix and intra-peptide $\beta$-sheet structures and their free energy differences from the global minimum state of the random coil structure are listed in Table 1. Note the summation of the three probabilities in Table 1 is slightly larger than 100%, because some residues have more than two kinds of secondary structures at the same time. Random coil is the major structure which has 64.9%. This probability is higher than the helix (29.4%) and $\beta$-sheet (6.3%) structures.

Free energy landscape of the two polyQ-peptide system as a function of the number of inter-peptide $\beta$-bridges and the length of helix calculated by eq. (12) is shown in Fig. 3. Some representative states were chosen and labeled from states A to N in Fig. 3. For the purpose of understanding the secondary structure changes and distinguishing the parallel and anti-parallel $\beta$-sheet, we chose the length of helix and number of $\beta$-bridges as the reaction coordinates here again. Figure 4 shows typical conformations and the transformation pathways between different states corresponding to Fig. 3. The relative positions of states A to N in Fig. 4 are similar to those in Fig. 3. The transformation pathways in Fig. 4 are constructed for neighboring states. State A has four parallel $\beta$-bridges. When the number of $\beta$-bridges decreases to two, the conformations becomes state B, which has shorter $\beta$-strands than state A. When the $3_{10}$-helix structure is formed additionally, state A transforms to the state G, and state B transform to state G or H. State C is the global minimum state in this
free-energy landscape. This state has a random coil structure, which can be reached directly from states B, D, and I. Through states D and E, state C can transform to state F. State F has seven $\beta$-bridges, which is the longest $\beta$-sheet conformation in our two polyQ-peptide system. The $\beta$-sheets sometimes coexist with the helix structures, as shown in the snapshots of states J and K. State J and K can be approached from state D, E, and I. As the length of the helix structure increases, the conformation changes into state L, at which there are four helix residues in one-peptide, into states M and N, at which both two polyQ peptides have the helix structures. As illustrated in Fig. 3, the anti-parallel $\beta$-sheet is more stable than the parallel. It is easier to form fully-extended anti-parallel $\beta$-sheet structures than the parallel $\beta$-sheet structures. When the length of helix becomes longer, the number of $\beta$-bridges gradually decreases and finally become zero when the length of helix is more than seven.

Besides helix and inter-peptide $\beta$-sheet structures, sometimes intra-peptide $\beta$-sheet structures appear during simulations. The probabilities of the three secondary structures are shown in Table 2. The summation of the four probabilities in Table 2 is larger than 100%, again. Comparing to table 1, there were more residues which have more than two kinds of secondary structures in the two polyQ-peptide system. These conformations can be found at states G, H, J, and K in Fig. 4. The random-coil state seems to be the global minimum in Fig. 4. However, if we consider all inter-peptide $\beta$-sheet conformations with different number of $\beta$-bridges, the total probability of inter-peptide $\beta$-sheet structures is larger than the random-coil structures, as shown in Table 2. The free energy difference was calculated from the global minimum state of the inter-peptide $\beta$-sheet structure. The fact that the inter-peptide $\beta$-sheet structure has the highest probability agrees well with the previous researches$^{9,25,28,37}$.

Comparing the ratio of the helix to the intra-$\beta$-sheet structure in the one and two polyQ-peptide systems, the proportion of the intra-$\beta$-sheet is relatively higher in the latter. We propose that the formation of the inter-peptide $\beta$-sheet in the two polyQ-peptide system begins when the intra-peptide $\beta$-sheet is formed and the helix structure is deformed.

To check our conjecture, we plot the inter-peptide distance against the total length of secondary structures of two PolyQ peptides in Fig. 5. The inter-peptide distance of two
PolyQ peptides is the separation between COM (center of mass) of Cα atoms of the first peptide and COM of Cα atoms of the second peptide. The production run was divided into three blocks, and the error bars of the interatomic distance was calculated as the standard deviation of the block averages. The inter-peptide distance increases and fluctuates when 10 or more residues form the helix structure. When only short helix structure is formed, the inter-peptide β-sheet is often formed as well, and the two peptides are close to each other. On the other hand, when a long helix structure is formed, there is no hydrogen bond between the two peptides, and the two peptides are sometimes close to each other, and sometimes separated. This is why the inter-peptide distance fluctuates when long helix structures are formed. The inter-peptide β-sheet is formed when the inter-peptide distance is less than about 6Å. The length of the inter-peptide β-sheet grows along with slight decrease of the inter-peptide distance. The intra-peptide β-sheet structures appear only when the inter-peptide distance is between 8 and 10Å. The inter-peptide distance of the intra-peptide β-sheet structures are between the inter-peptide distances of helices and inter-peptide β-sheets. From these results, we can have the following picture on dimer formation of the polyQ peptides: The polyQ peptides tend to form the helix structure rather than the β-sheet structure when they are separated, as in the one polyQ peptide system. They transform to inter-peptide β-sheets when the inter-peptide distance becomes shorter. The intra-peptide β-sheet structures can be found between the helix to inter-peptide β-sheet transformation.

To clarify further the relation between the inter-peptide distance and the peptide structures, we used eq. (12) to calculate the free energy landscape as a function of inter-peptide distance and end-to-end distance and showed the results in Fig. 6. The end-to-end distance of two polyQ peptides is defined as the average of two distances between the first and the last Cα atoms in each of two polyQ peptides. Global- and local-minimum free-energy states were labeled from A to J, as shown in Fig. 6; such notations have nothing to do with those in Fig. 3. Typical conformations and the transformation pathways from states A to J are illustrated in Fig. 7. The β-sheet structure is formed at state A, which has the shortest inter-peptide distance and the longest end-to-end distance. When the two polyQ peptides get slightly separated and more twisted, state A is transformed to state B, C, and F sequentially. State B is the global-minimum state on this free energy landscape shown in Fig. 6.
It forms both helix and β-sheet as shown in Fig. 7. The intra-peptide β-sheet is seen at state F as an intermediate state between the helix and β-sheet structures. However, this intra-peptide β-sheet is not an obligatory intermediate state, because helix structure can transform to β-sheet structure without intra-peptide β-sheet through the pathway E → B → A or E → D → A. Through the pathway E → B → A, we can say that random coil structure is not necessary in the helix-sheet transformation, too. Following the conformations at states F, C, B, and A, we can find that the polyQ peptides extend when they get close to each other. State A, B, C, and F can be obtained from state D and E, the free energies of which are almost the same as state F. Structures can be changed further between state D, E, F, and G. State G can be approached from states H and I, which are close to state J. Similar to Fig. 6, we can see that the peptides does not form inter-peptide β-sheet when the inter-peptide distance is too long, as shown in snapshots D to J in Fig. 7.

The global minimum in Fig. 3 is different from that in Fig. 6. This is because the reaction coordinates are different between these free-energy landscape. The reaction coordinates $\xi_1$ and $\xi_2$ in Fig. 3 are the length of helix and number of β-bridges, respectively. The helix and β-strand structures were determined by DSSP. DSSP is known as a relatively strict criteria for the secondary structure determination, because it requires hydrogen bonds between backbone residues (e.g. see ref. 48). On the other hand, $\xi_1$ and $\xi_2$ in Fig. 6 are the inter-peptide distance and the end-to-end distance, respectively. In Fig. 3 state C seems the global-minimum free-energy state, which corresponds to the random coil structure. Because the inter-peptide distance and the end-to-end distance of the random coil structure are usually large and widely distributed, the random coil structure can be found in a broad range of Fig. 6. For example, Fig. 8 shows other conformations that were observed at states A and E in Fig. 6. Although most of the conformations at states A form β-strand as in Fig. 7, there are a few random coil conformations. They are also extended but do not form a β-strand, as in Fig. 8. Similarly at state E, there are some conformations that do not have a helix structure, but only a turn structure, as shown in Fig. 8. These conformations are similar to those in Fig. 7, but have no secondary structures. This is because some hydrogen bonds are not recognized by DSSP if hydrogen or oxygen positions are slightly different. It means that the random coil structure, which is observed at the global minimum state C in Fig. 4, includes many
types of conformations. One conformation is close to the $\alpha$-helix and another conformation is close to the $\beta$-sheets. Thus, the random coil structure cannot be observed as the global-minimum free-energy state in Fig. 6. As in these figures, the global-minimum free-energy state depends on the reaction coordinates in general\textsuperscript{48}.

In our case, we focused on the transformation pathways between the $\alpha$-helix and $\beta$-sheet structures. As we can see in Fig. 4, a random coil structure can be seen between the $\alpha$-helix and $\beta$-sheet. In Fig. 3, we further distinguish the parallel and anti-parallel $\beta$-sheet and show that the anti-parallel $\beta$-sheet is more dominant. On the other hand, we focused on how the conformation changes as the two peptides get closer in Figs. 6 and 7. The $\alpha$-helix structure is often seen when they are separated. As they get closer, the inter-peptide $\beta$-sheets are formed. A random coil structure can be often seen at states C and D between the $\alpha$-helix and $\beta$-sheet, again in Fig. 7. In this way, the conformations in Figs. 4 and 7 show different aspects of the transformation pathways. Thus, both free energy landscapes in Figs. 4 and 7 provide useful information.

**CONCLUSIONS**

The polyQ fragment in a protein can cause amyloidoses\textsuperscript{16–20}. It is important to understand how the polyQ peptides aggregate and form an amyloid fibril. To understand the aggregation mechanism, we performed 100 ns all-atom REMD simulations for one and two polyQ peptides with explicit water molecules.

From simulation data, we find that single (or separated) polyQ peptide tends to form random coils, but sometimes helix or intra-peptide $\beta$-sheet structures. When the inter-peptide distance decreases, the helix structure is unfolded, and the intra-peptide $\beta$-sheet structure is formed with a higher probability. The intra-peptide $\beta$-sheet was also found as an free-energy local-minimum state, that is an intermediate state between helix structure and inter-peptide $\beta$-sheet. As the inter-peptide distance decreases further, the polyQ peptides extend and form the inter-peptide $\beta$-sheets. The inter-peptide $\beta$-sheet structure is the main structure in the two polyQ-peptide system.

Some conflicting models have been proposed so far on the anti-parallel $\beta$-sheet structure\textsuperscript{9}.
or parallel \( \beta \)-sheet structure\(^{25,26}\). Our detailed molecular dynamics simulations showed that inter-peptide \( \beta \)-sheet was the main structure in the two polyQ-peptide system. In the inter-peptide \( \beta \)-sheet, the anti-parallel \( \beta \)-sheet is more stable than parallel \( \beta \)-sheet. Our result is consistent with the model previously proposed from experiments\(^{24,33}\) and a coarse-grained molecular dynamics simulation\(^9\).

**ACKNOWLEDGMENTS**

We thank M. S. Li for useful discussion. This work was supported by NSC 100-2911-I-001-507, NSC 100-2923-M-001-003-MY3, NSC 102-2112-M-001-006, NCTS (NORTH), and International Collaboration Program of Institute for Molecular Science. The computations were performed on the computers at the Research Center for Computational Science, Okazaki Research Facilities, and National Institutes of Natural Sciences.

**References**


Figure 1: Free energy landscape of one polyQ peptide at $T = 300$ K as a function of the length of helix and the number of $\beta$-bridges. Some letters are colored in white to make them read easily in contrast of the background color.

Figure 2: Conformations and the transformation pathways of the global- and local-minimum free-energy states A to I of the system one polyQ-peptide and 1611 water molecules. The peptide conformations were created using RasMol$^{59}$.

Figure 3: Free energy landscape of two polyQ peptides at $T = 300$ K as a function of the length of helix and the number of $\beta$-bridges. Some letters are colored in white to make them read easily in contrast of the background color.

Figure 4: Conformations and the transformation pathways of the global- and local-minimum free-energy states A to N of the system with two polyQ-peptides and 3129 water molecules. The figures were created using RasMol$^{59}$.

Figure 5: Relation between the inter-peptide distance and the total length of helix, inter-peptide $\beta$-sheet, and intra-peptide $\beta$-sheet of the two polyQ-peptide system.

Figure 6: Free energy landscape of the two polyQ-peptides as a function of the inter-peptide distance and the end-to-end distance. Some letters are colored in white to make them read easily in contrast of the background color.

Figure 7: Conformations and the transformation pathways of the global and local minimum states on the free energy landscape of the inter-peptide distance and the end-to-end distance. The figures were created using RasMol$^{59}$.

Figure 8: Random coil conformations observed at states A and E. The figures were created using RasMol$^{59}$.
Figure 1
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu
Figure 2
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu
Figure 3
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu
Figure 4
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu
Figure 5
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu
Figure 6
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu
Figure 7
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu

Figure 8
Hsin-Lin Chiang, Chun-Jung Chen, Hisashi Okumura, Chin-Kun Hu
<table>
<thead>
<tr>
<th>structure</th>
<th>probability(%)</th>
<th>$\Delta F$(kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>helix</td>
<td>29.4±4.6</td>
<td>0.47±0.12</td>
</tr>
<tr>
<td>intra-β-sheet</td>
<td>6.3±5.1</td>
<td>1.39±0.49</td>
</tr>
<tr>
<td>random-coil</td>
<td>64.9±8.9</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1: Probability and free energy difference of each secondary structure in the one polyQ-peptide system.

<table>
<thead>
<tr>
<th>structure</th>
<th>probability(%)</th>
<th>$\Delta F$(kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>helix</td>
<td>23.1±3.2</td>
<td>0.47±0.15</td>
</tr>
<tr>
<td>inter-β-sheet</td>
<td>50.5±10.2</td>
<td>-</td>
</tr>
<tr>
<td>intra-β-sheet</td>
<td>11.8±5.7</td>
<td>0.87±0.31</td>
</tr>
<tr>
<td>random-coil</td>
<td>25.3±10.7</td>
<td>0.41±0.28</td>
</tr>
</tbody>
</table>

Table 2: Probability and free energy difference of each secondary structure in the two polyQ-peptide system.