J Mol Model (2013) 19:1627–1639
DOI 10.1007/s00894-012-1712-9

ORIGINAL PAPER

Study of the aggregation mechanism of polyglutamine peptides using replica exchange molecular dynamics simulations

Miki Nakano · Kuniyoshi Ebina · Shigenori Tanaka

Received: 17 May 2012 / Accepted: 27 November 2012 / Published online: 5 January 2013

Abstract Polyglutamine (polyQ, a peptide) with an abnormal repeat length is the causative agent of polyQ diseases, such as Huntington's disease. Although glutamine is a polar residue, polyQ peptides form insoluble aggregates in water, and the mechanism for this aggregation is still unclear. To elucidate the detailed mechanism for the nucleation and aggregation of polyQ peptides, replica exchange molecular dynamics simulations were performed for monomers and dimers of polyQ peptides with several chain lengths. Furthermore, to determine how the aggregation mechanism of polyQ differs from those of other peptides, we compared the results for polyQ with those of polyasparagine and polyleucine. The energy barrier between the monomeric and dimeric states of polyQ was found to be relatively low, and it was observed that polyQ dimers strongly favor the formation of antiparallel β-sheet structures. We also found a characteristic behavior of the monomeric polyQ peptide: a turn at the eighth residue is always present, even when the chain length is varied. We previously showed that a structure including more than two sets of β-turns is stable, so a long monomeric polyQ chain can act as an aggregation nucleus by forming several pairs of antiparallel β-sheet structures within a single chain. Since the aggregation of polyQ peptides has some features in common with an amyloid fibril, our results shed light on the mechanism for the aggregation of polyQ peptides as well as the mechanism for the formation of general amyloid fibrils, which cause the onset of amyloid diseases.

Keywords Polyglutamine disease · Protein aggregation · Protein folding · Free-energy landscape · Replica exchange molecular dynamics simulation · Amyloid fibril

Abbreviations

Q	Glutamine
polyQ	Polyglutamine
MD	Molecular dynamics
REMD	Replica exchange molecular dynamics
MMD	Multiple molecular dynamics
CD	Circular dichroism
DPSS	Dictionary of protein secondary structure
apβ	Antiparallel β-sheet
pβ	Parallel β-sheet
$D_{\min}$	The minimum distance between the C_α atoms of the monomers (except for both termini)
Λ_{mono}	Distance from the C_α atom of the C-terminus to that of the N-terminus
$n_{\mathrm{Hbond}}^{\mathrm{inter}}$	The number of intermolecular hydrogen bonds between two molecules

Introduction

Polyglutamine (polyQ) diseases are neurodegenerative disorders, including Huntington's disease. These diseases are caused by an abnormal expansion of the CAG repeat coding the polyQ region that appears in the corresponding pathogenesis protein [1–3]. It is a common feature of all polyQ diseases that if the length of the CAG repeat in each pathogenesis protein becomes longer than the critical length, these proteins will misfold and aggregate to form insoluble,

M. Nakano (✉) · K. Ebina
Graduate School of Human Development and Environment,
Kobe University, 3-11, Tsurukabuto, Nada,
Kobe 657-8501, Japan
e-mail: nakano@eniac.scitec.kobe-u.ac.jp

S. Tanaka (✉)
Graduate School of System Informatics, Department of
Computational Science, Kobe University, 1-1, Rokkodai, Nada,
Kobe 657-8501, Japan
e-mail: tanaka2@kobe-u.ac.jp

amyloid-like fibrils, which cause neuronal dysfunction and eventual cell death [4–9].

The polyQ peptide is a very specific peptide that consists of only glutamine (Q) residues, and its properties vary significantly with its chain length. However, this peptide has some features in common with other amyloid-forming peptides—it has a simple sequence, a low hydrophobic content, and several metastable structures. The aggregation process is an important step in the onset of polyQ diseases, and such aggregation processes are also an important feature of amyloid diseases. Therefore, elucidating the mechanism for the aggregation of polyQ peptides will not only yield useful information for the effective treatment of polyQ diseases but also provide insights into the pathogenic mechanisms of other amyloid diseases.

CAG-expanded polyQ peptides are known to form β-sheet structures, as shown by X-ray diffraction, circular dichroism (CD), Fourier transform infrared spectroscopy, and other experimental methods [10–14]. However, recent studies indicate that monomeric polyQ has no specific secondary structure. Chen et al. [5] measured the CD spectrum of $K_2Q_{42}K_2$, demonstrating that polyQ peptides transform into β-sheet structures from the random-coil state during aggregation. By performing multiple molecular dynamics (MMD) simulations for short peptides, Vitalis et al. [15] found that the ensemble structure of polyQ is disordered. Using fluorescence correlation spectroscopy measurements, Crick et al. [16] suggested that the ensemble of monomeric polyQ should consist of a heterogeneous collection of collapsed structures. Further, Masino et al. [17] characterized the structure of polyQ in a solution as a random coil using a model system based on glutathione *S*-transferase fusion proteins. In our previous study [18], we revealed that a monomeric polyQ peptide frequently forms a β-turn structure in a certain fraction, and that this peptide dynamically changes between distinct structures. Furthermore, we showed that these β-turn structures are stabilized by the formation of oligomers owing to the increase in the number of hydrogen bonds between the main chains.

As mentioned above, the structural properties of the polyQ monomer and its aggregates have been revealed. Furthermore, the growth process of polyQ aggregation is known to involve a seeding effect analogous to general crystal-growth processes [3, 5, 9]. However, unsolved questions remain: what are the dynamical processes that occur from nucleation to aggregation of polyQ peptides, and how large is the binding energy of polyQ peptides in water? In addition, why does the polyQ peptide have a high tendency to aggregate compared with other amino acids?

To answer these questions, we need to investigate the precise features of the free-energy landscape of polyQ in water. For the aggregation process of polypeptides, it is expected that the internal structural changes of each peptide play an important role. Therefore, the energy landscape should be calculated with an all-atom model. In addition, to investigate the specificity of the polyQ peptide, the structural properties of polyQ peptides should be examined and compared with those of other homopolyamino-acid peptides. Some molecular dynamics (MD) simulations of long polyQ chains have been performed with coarse-grained models [19–21]. In the study described in the present paper, we performed detailed investigations of the properties of all-atom models of relatively short polyQ peptides to clarify the characteristics of the glutamine residues.

Since the free-energy landscape of the polyQ dimer is poorly resolved, we needed to explore a wide range of energy space of the polyQ peptides. As a result, we faced additional challenges: since the appearance ratio of the state of interest is low, a sufficient number of samples is required to investigate the properties of this state, and all-atom simulations usually also require a significant amount of calculation time. On the other hand, standard MD configurations are often trapped within local energy minima.

To overcome these difficulties, we used replica exchange molecular dynamics (REMD) simulations [22] with an all-atom model. This technique can explore a wide range of energy space without becoming trapped in local energy minima within an attainable simulation time. Multiple molecular dynamics (MMD) simulation is also useful for effective sampling. In this technique, we executed several simulations starting from the same initial structure but with different sets of initial velocities. This method allowed us to perform effective sampling in the conformational space of proteins from multiple short trajectories rather than from a single long trajectory [23–25].

In the present study, we used an implicit water solvent model to explore the nature of polyQ peptides caused by the interactions between side chains. The hydrogen-bond network between peptide and water may play an important role in the aggregation process of polyQ. However, we first investigated the interactions between peptides as the primary problem. Because the computational cost of using an implicit water model was relatively low, we were able to analyze a sufficient amount of data to observe the essential features caused by the interactions between peptides. In addition, the computational time is approximately proportional to the square of the number of atoms included in the simulation model. In order to obtain a sufficient number of samples, we chose a strategy that involved calculating relatively short peptides rather than long peptides.

By analyzing these results from the simulations for several lengths of short peptides, a richer amount of information should be obtained than that gained from a single MD simulation performed for a long peptide. In order to elucidate how the aggregation mechanism of polyQ differs from those of other peptides, we also compared the results for

polyQ with those of polyasparagine (polyN) and polyleucine (polyL). Leucine is a typical hydrophobic residue with a similar molecular size to glutamine, whereas asparagine is chemically similar to glutamine except that the side chain is shorter by one methylene group. Combining these results, we attempted to clarify the role of side chains and the specificities of the aggregation mechanism of polyQ peptides.

This paper is organized as follows. In the "Materials and methods" section, we describe the materials and methods that we used. In the "Results" section, we show the results of our simulations, and these results are discussed in the "Discussion" section. Finally, we draw conclusions based on the results of this study in the "Conclusions" section.

Materials and methods

Materials

In this study, we prepared three types of models for our molecular dynamics (MD) simulations. The first type consisted of those with several chain lengths of monomeric polyQ peptides: Q_9, Q_{11}, Q_{13}, Q_{15}, and Q_{18}. The chain lengths of these monomers were 9, 11, 13, 15, and 18 in units of Q, respectively. All of the initial structures of these polyQ monomers were extended. The second type of model represented pairs of polyQ monomers with several chain lengths: Q_3, Q_5, Q_7, and Q_9. In these cases, we set the initial structures as follows: the directions of two monomer chains were parallel with each other, and the distances *R* between the main chains of each monomer were 10 Å. The third type of model comprised Q_7 dimers in which the initial distances *R* between the main chains of each monomer were 5 Å. In this third type of model, the main chains were set to be either antiparallel or parallel to each other. Figure 1 shows the initial structures for each model.

Furthermore, we constructed models consisting of either leucine or asparagine for both monomers and pairs. Both asparagine and leucine have similar structures to glutamine. Leucine is a typical hydrophobic residue with a similar size to glutamine, and it has no phenyl groups nor sulfur atoms. On the other hand, asparagine is a very similar residue to glutamine—they both have an amide group at the end of their side chain. The sole difference between glutamine and asparagine is the length of this side chain. Comparing the results of the simulations for these peptides, we attempted to specify the character of the polyQ peptide.

All peptides were constructed using the LEaP module included in the Amber10 package [26]. The N- and C-termini were capped with acetyl and *N*-methyl groups, respectively. The initial structures of all of these samples were fully extended. The list of peptides that we used in these REMD simulations is shown in Table 1.

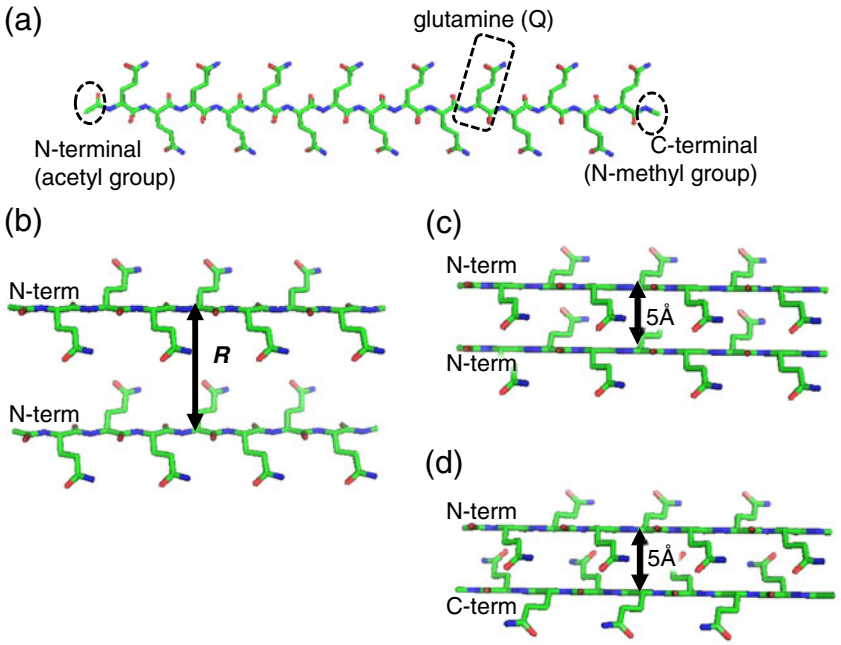

Fig. 1a–d Initial structures utilized in our molecular dynamics simulations. **a** Initial structure of a Q_{15} monomer employed for REMD simulation. **b** Initial structure of a Q_7 pair used for REMD simulation. *R* is the distance between the central C_α atoms of each monomer. The initial value of *R* was set to be 10 Å, and it was restricted to <40 Å during the REMD simulations. **c**,**d** Initial structures of Q_7 dimers which were used for ordinary molecular dynamics simulations. The distances between the central C_α atoms of each monomer were set to 5 Å, and no constraints were applied. Carbon atoms are *green*, oxygen atoms are *red*, and nitrogen atoms are *blue*. Hydrogen atoms are omitted for clarity. These figures were created with PyMOL [38]

Table 1 List of the peptide names, the number of atoms, number of replicas, and the temperatures that were used in the REMD simulations for the monomer models and pair models

	Peptide name	Number of atoms	Number of replicas	Temperatures for each replica
Monomer models	Q_9	165	14	252.1, 275.0, 300.0, 327.2, 356.9, 389.3, 424.7, 463.3, 505.3, 551.2, 601.3, 655.9, 715.4, 780.4
	Q_{11}	199	14	277.2, 300.0, 324.7, 351.4, 380.3, 411.6, 445.4, 482.1, 521.8, 564.7, 611.1, 661.4, 715.8, 774.7
	Q_{13}	233	16	259.2, 278.8, 300.0, 322.7, 347.1, 373.4, 401.7, 432.2, 464.9, 500.1, 538.0, 578.8, 622.6, 669.8, 720.5, 775.1
	Q_{15}	267	16	280.2, 300.0, 321.1, 343.8, 368.0, 394.0, 421.8, 451.6, 483.4, 517.5, 554.1, 593.1, 635.0, 679.8, 727.8, 779.1
	Q_{18}	318	18	264.8, 281.9, 300.0, 319.3, 339.9, 361.8, 385.1, 409.9, 436.3, 464.4, 494.4, 526.2, 560.1, 596.2, 634.6, 675.5, 719.0, 765.3
	N_{15}	222	14	278.4, 300.0, 323.3, 348.4, 375.5, 404.7, 436.1, 470.0, 506.5, 545.8, 588.2, 633.9, 683.2, 736.3
	L_{15}	297	18	263.6, 281.2, 300.0, 320.0, 341.4, 364.2, 388.5, 414.4, 442.1, 471.6, 503.1, 536.6, 572.4, 610.7, 651.4, 694.9, 741.3, 790.8
Pair models	Q_3 pair	126	12	245.8, 271.5, 300.0, 331.4, 366.1, 404.5, 446.8, 493.6, 545.4, 602.5, 665.6, 735.3
	Q_5 pair	194	14	276.9, 300.0, 325.0, 352.1, 381.5, 413.3, 447.7, 485.1, 525.5, 569.3, 616.8, 668.2, 723.9, 784.3
	Q_7 pair	262	16	280.0, 300.0, 321.3, 344.2, 368.8, 395.0, 423.2, 453.3, 485.6, 520.3, 557.3, 597.0, 639.6, 685.2, 734.0, 786.3
	Q_9 pair	330	18	265.4, 282.2, 300.0, 319.0, 339.1, 360.6, 383.3, 407.6, 433.3, 460.7, 489.8, 520.8, 553.7, 588.7, 625.9, 665.4, 707.5, 752.2
	N_3 pair	106	12	241.3, 269.1, 300.0, 334.5, 372.9, 415.8, 463.6, 516.8, 576.2, 642.5, 716.3, 798.6
	N_5 pair	164	12	275.0, 300.0, 327.3, 357.1, 389.7, 425.2, 463.9, 506.2, 552.3, 602.6, 657.5, 717.4
	N_7 pair	220	14	278.3, 300.0, 323.4, 348.7, 375.9, 405.2, 436.9, 471.0, 507.7, 547.3, 590.1, 636.1, 685.8, 739.3
	N_9 pair	276	16	280.5, 300.0, 320.8, 343.0, 366.8, 392.3. 419.5, 448.6, 479.7, 512.9, 548.5, 586.5, 627.2, 670.7, 717.2, 766.9
	L_7 pair	290	18	263.2, 281.0, 300.0, 320.2, 341.9, 365.0, 389.7, 416.0, 444.1, 474.1, 506.2, 540.4, 576.9, 615.9, 657.5, 701.9, 749.4, 800.0

Replica exchange molecular dynamics (REMD)

Generally, in conventional MD simulations, the system often becomes trapped in local potential energy minima. In order to explore a wide region of conformational space without becoming trapped in local energy minima, we performed REMD simulations [22].

In REMD, the replica in each canonical ensemble at a fixed temperature is simulated independently for a certain number of MD steps. Pairs of replicas at neighboring temperatures are then exchanged with the acceptance probability

$$w(x_i|x_j) = \left\{ \begin{array}{ll} 1 & ;\text{for } \Delta \le 0 \\ \exp(-\Delta) & ;\text{for } \Delta > 0 \end{array} \right\}. \quad (1)$$

Here, $\Delta=(1/k_BT_i-1/k_BT_j)(E_i-E_j)$ and k_B is the Boltzmann constant. Subscripts i and j refer to the replica numbers; x_i and x_j are the sets of coordinates of all the atoms at absolute temperatures T_i and T_j, with energies of E_i and E_j, respectively. After the exchange, the simulations resume at the new temperatures.

REMD conditions

For REMD simulations, we used the REMD module of Amber10 [26] in combination with the parm96 force field and the generalized Born implicit solvent model of Onufriev, Bashford, and Case [27]. This combination of force field and solvation model provides reasonable structures of proteins and a fair balance between α-helices and β-sheets [28, 29]. We set the number of replicas to be the square root of the number of atoms included in each model [30]. The number of atoms, number of replicas, and the temperatures used for each model are listed in Table 1.

Because we did not use periodic boundary conditions in this study, molecules would wander through infinite space if no constraints were applied; once two monomers separate from each other, they would seldom approach again. Therefore, we performed REMD simulations for the pair models with a distance constraint, R, applied between the monomers, in the following manner. We set the constraint conditions using the restraint function $V_{res}(R)$, where R is the distance between the central C_α atoms of each monomer (see Fig. 1):

$$V_{res}(R) = \left\{ \begin{array}{ll} 0 & ;\text{for } R \le r_1 \\ k(R-r_1)^2 & ;\text{for } r_1 \le R < r_2 \\ \lambda(R-r_2) + V_{res}(r_2) & ;\text{for } r_2 \le R \end{array} \right\}. \quad (2)$$

Here, λ is the slope of the parabola at the point $R=r_2$, and the value is determined by the force constant k. We set these parameters to the following values: r_1=40 Å, r_2=50 Å, and

k=10 kcal/mol/Å^2. The constraint length, r_1=40 Å, was much longer than the length of the molecule.

Energy minimizations were performed to relax the initial structures for each model. Minimization was carried out by implementing 1000 steps of the conjugate gradient method, followed by 500 steps of steepest descent calculations, using the multisander module included in Amber10. Before initiating replica exchange, we performed 200 ps simulations with a time step of 2 fs to ensure that each replica was equilibrated at its individual temperature. All replicas were prepared in the same structure, the energy of which was minimized using the method described above. Subsequently, the peptides were heated to their respective temperatures using the Langevin thermostat [31, 32] with a collision frequency γ of 1.0 ps^{-1}. After equilibration, the REMD simulations were started.

All simulations were performed with a 2 fs time step, a replica exchange interval of 1 ps, a nonbonded cutoff length of 99 Å, and an external dielectric constant, ε, equal to 78. Langevin dynamics were applied to control the temperature, with a collision frequency γ of 1.0 ps^{-1}. Throughout the simulations, the lengths of bonds involving hydrogen atoms were fixed using the SHAKE algorithm. The corresponding acceptance ratios for replica exchange were 15–20 %.

We defined $D_{\min}$ as the minimum distance between the C_α atoms of the monomers (except for both termini). Figure 2 shows the accumulated averages of $D_{\min}$ ($\overline{D_{\min}}$) for each pair model at 300 K. We analyzed the snapshots at 300 K during the period after the $\overline{D_{\min}}$ values had been sufficiently equilibrated. The sampling times that we used in our analysis and the number of snapshots for each model are listed in Table 2. All samples were taken from the replicas at 300 K every 20 ps during the sampling time intervals listed in Table 2.

Table 2 List of the sampling time and number of samples for each model

	Peptide name	Sampling time (ns)[a]	Number of samples[b]	Number of dimerized samples[c]
Monomer models	Q_9	50–300	12 500	–
	Q_{11}	50–300	12 500	–
	Q_{13}	50–300	12 500	–
	Q_{15}	50–300	12 500	–
	Q_{18}	150–500	17 500	–
	N_{15}	50–300	12 500	–
	L_{15}	50–300	12 500	–
Pair models	Q_3 pair	150–500	17 500	44
	Q_5 pair	150–500	17 500	171
	Q_7 pair	500–1000	25 000	811
	Q_9 pair	500–1000	25 000	149
	N_3 pair	150–500	17 500	67
	N_5 pair	150–500	17 500	197
	N_7 pair	150–500	17 500	223
	N_9 pair	150–500	17 500	81
	L_7 pair	150–500	17 500	3206

[a] These times show the MD periods during which sampling was performed for each replica. For example, 50–300 indicates that replica sampling started at t=50 ns and ended at t=300 ns after a thermalization. The sampling rate was once per 20 ps for all peptides.

[b] The number of samples that were taken during the analysis for each model. These snapshots were taken from the replicas at 300 K during the sampling time intervals shown.

[c] The number of dimerized samples for which the minimum distance ($D_{\min}$) between monomers was shorter than 6 Å at 300 K

Multiple molecular dynamics (MMD) simulation

To investigate the more stable polyQ dimer structure, we also performed MMD simulations for two types of Q_7

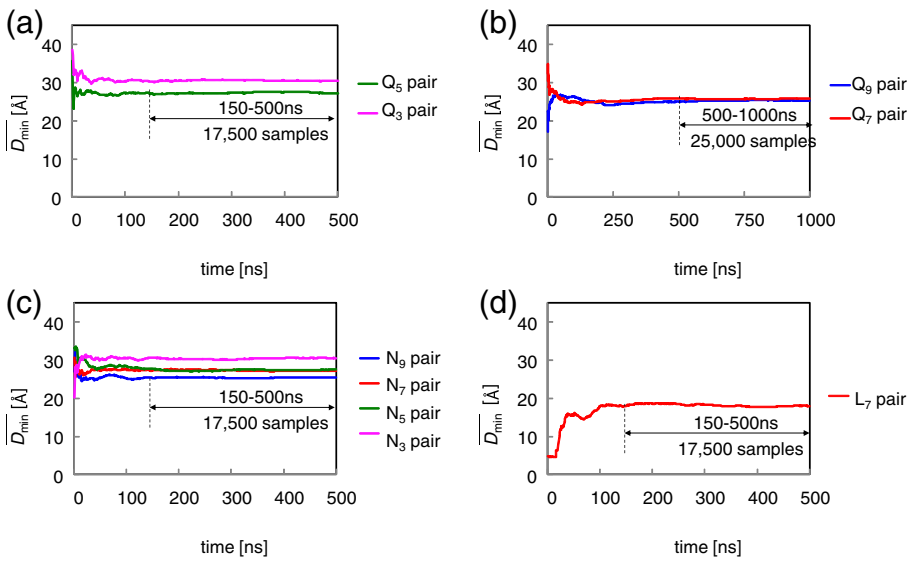

Fig. 2 Accumulated average of the minimum distance between monomers ($\overline{D_{\min}}$) for each pair model. We defined $D_{\min}$ as the minimum distance between the C_α atoms of the monomers (except for both termini), and $\overline{D_{\min}}$ as the temporally accumulated average of $D_{\min}$. *Arrows and times* in these figures show the MD periods during which sampling was performed for each model. The number of samples for each model is also shown

dimers; in one type the main chains were parallel, whereas they were antiparallel in the other type. Sampling of multiple short trajectories with different initial velocities was undertaken to improve the efficiency of conformational sampling in the MD simulation of the polypeptides. This approach has been shown to be more efficient than the use of a single long trajectory [23–25]. Therefore, we performed ten MD simulations with different random seeds for the initial velocities from both the antiparallel and parallel structures described above, respectively. These simulations were carried out at 300 K. All other conditions were the same as those of REMD.

Results

REMD for monomer models

Figure 3a shows the distributions of the secondary structures of the Q_{15}, N_{15}, and L_{15} monomers. The definitions of the classifications for each residue in a monomer are listed in Table 3. The Dictionary of Protein Secondary Structure (DPSS) software [33] was used to determine the secondary structure of each residue. The N_{15} and L_{15} peptides predominantly adopted α-helical structures, whereas Q_{15} adopted various structures, including the β-turn.

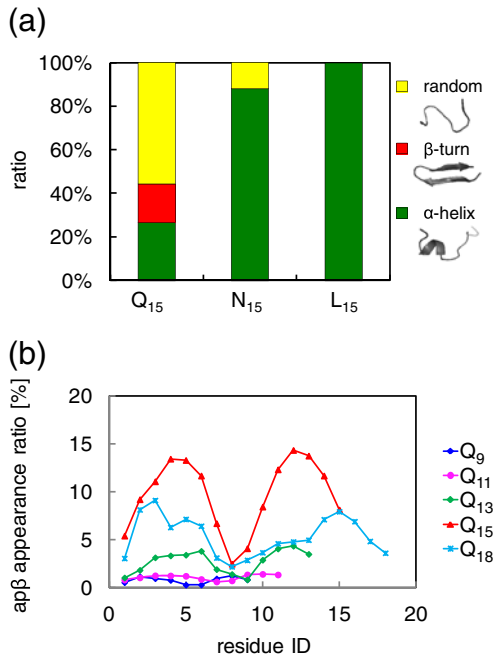

Fig. 3 **a** Distribution of secondary structures at 300 K for Q_{15}, N_{15}, and L_{15}, as calculated in REMD simulations. Definitions of these structures are provided in Table 2. *Green bar*: α-helix. *Red bar*: β-turn. *Yellow bar*: random coil. Representative structures for each secondary structure that we defined are also shown. **b** Appearance ratios of snapshots with antiparallel β-strand (*apβ*) properties at each residue for different polyQ monomer lengths

Figure 3b shows the frequency with which an antiparallel β-strand appeared at each residue (the apβ appearance ratio) for polyQ monomers of different lengths. As the length of the polyQ chain increased, from Q_9 to Q_{15}, the apβ appearance ratio increased, whereas the ratios for Q_{18} were slightly decreased compared to those of Q_{15}. Interestingly, the β-strand always turned at the same site (the eighth site), regardless of the chain length.

REMD for pair models

Figure 4 shows the normalized distribution function $p(D_{\min})$ of the minimum distance $D_{\min}$ between monomers for each pair model. $p(D_{\min})$ was calculated based on the appearance ratio from snapshots obtained from REMD simulations, using the following equation:

$$\int_0^{r_2} 4\pi {D_{\min}}^2 p(D_{\min}) dD_{\min} = \frac{4\pi {r_2}^3}{3}. \quad (3)$$

Here, r_2 is the parameter defined by Eq. 2, which was used to constrain the distance between monomers.

For polyQ or polyN models, sharp peaks exist in $p(D_{\min})$, and the centers of these peaks are located at around 4.5 Å. This observation indicates that the main chains of two monomers are connected by hydrogen bonds. In contrast, for L_7, a higher and broader peak appears than those observed for Q_7 and N_7. For polyL peptides, monomers are clustered together via hydrophobic interactions without bonding. We regarded a snapshot as a "dimerized" sample when the value of $D_{\min}$ was less than 6 Å. The number of dimerized samples for each model is listed in Table 2 for all pair models.

Figure 5 shows the free-energy landscapes at 300 K for the pair models of Q_7, N_7, and L_7, where we adopted the minimum distance $D_{\min}$ between monomers and the length of the monomer extension Λ_{mono} as two-dimensional reaction coordinates. Here, Λ_{mono} was defined as the distance from the C_α atom of the C-terminus to that of the N-terminus. We calculated the free energies in terms of the potential of the mean force:

$$G = G_0 + k_B T \ln(p/p_0). \quad (4)$$

Here, G is the free energy at the absolute temperature T, and p is the population of snapshots at the designated reaction coordinates. G_0 is the standard free energy calculated at the coordinate where the population of snapshots is p_0 (an arbitrary choice).

As seen in Fig. 5, when two monomers existed separately, they took various values of Λ_{mono} for Q_7 and N_7, whereas only small values of Λ_{mono} appeared for L_7. On the other

Table 3 Definitions of the classes of secondary structure used for the snapshots obtained from REMD simulations of Q_{15}, N_{15}, and L_{15} monomers. The program DPSS [33] was used to determine the secondary structure of each residue

Class	Definitions
α-Helix	Snapshot satisfies both of the following conditions:
	(1) There are ≥4 residues with α-helix properties
	(2) There are ≥2 hydrogen bonds between the main chains
β-Turn	Snapshot satisfies both of the following conditions:
	(1) There are ≥4 residues with antiparallel β properties
	(2) There are ≥1 hydrogen bonds between residues separated by more than 3 residues
Random	Snapshot simultaneously satisfies the definitions of an α-helix and a β-turn, or does not satisfy either

hand, when two monomers approached each other (i.e., when D_{min} was small), most of the Q_7 monomer samples became extended, whereas the N_7 and L_7 monomers took various values. Representative structures for each peptide are also shown in the inset of Fig. 5, along with their associated basins. When D_{min} was small, most of the Q_7 dimer structures were antiparallel β-sheets, whereas the N_7 and L_7 monomers took a variety of structures, including α-helices and β-sheets.

Figure 6 shows the projections onto the D_{min} and ΔG planes of the free-energy landscapes depicted in Fig. 5. We estimated the statistical errors by calculating the standard deviations among 100 data sets including 1000 samples which were selected at random. The figures for Q_7 and N_7 show energy barriers between the dimer and monomer states as low as 1.2–2 kcal/mol, which are thus easily overcome at physiological temperatures. The figure for L_7 shows that the

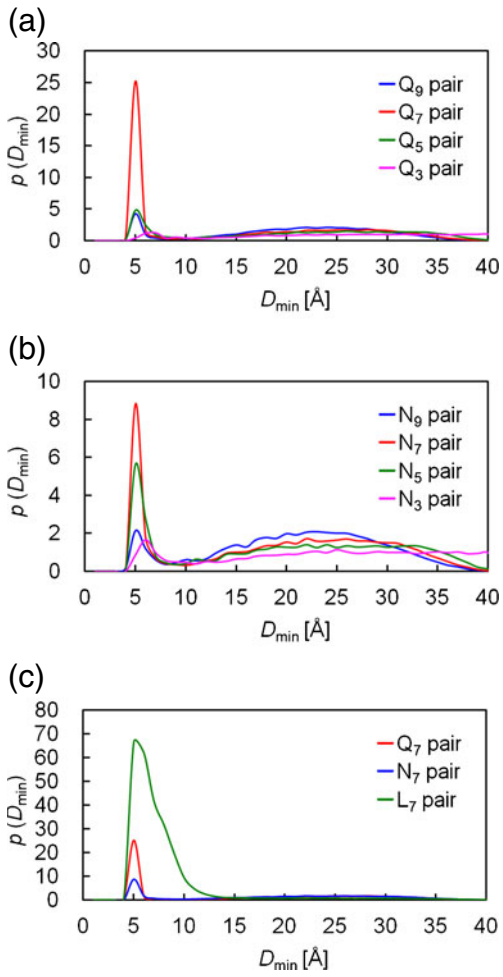

Fig. 4a–c Normalized distribution function $p(D_{min})$ of the minimum distance D_{min} between monomers for each pair model at 300 K, where $p(D_{min})$ is defined in Eq. 3 (see text). **a** Results for polyQ pairs. **b** Results for polyN pairs. **c** Results for the Q_7, N_7, and L_7 pairs

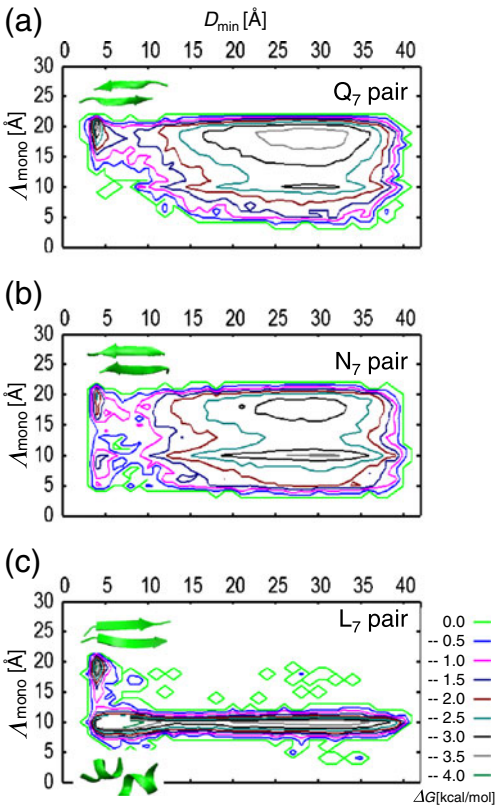

Fig. 5 Free-energy landscapes at 300 K for the Q_7, N_7, and L_7 pairs, which are expressed as a function of the minimum distance (D_{min}) between monomers and the length of the monomer extension (Λ_{mono}). *Arrows* connect representative dimerized structures with their associated basins. Free energies were calculated by the potential of mean force method using Eq. 4

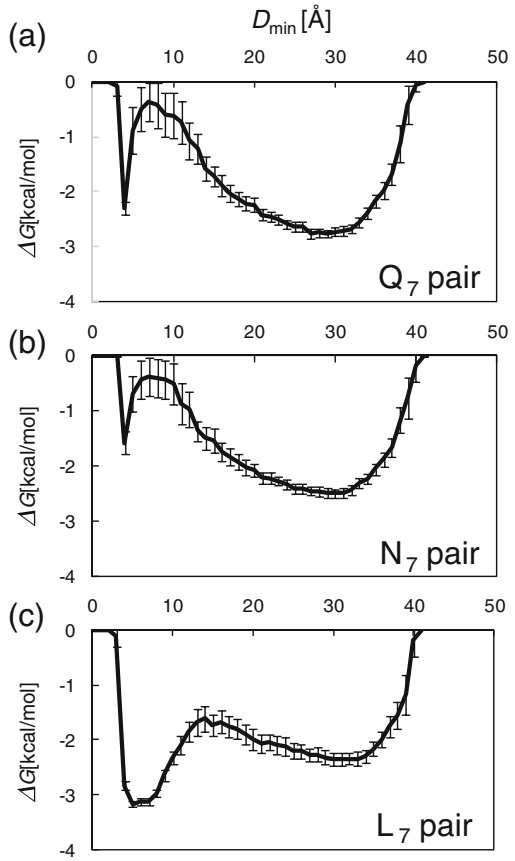

Fig. 6a–c Projections of the free-energy landscapes shown in Fig. 5 onto D_{min} and ΔG planes. Statistical errors were estimated by calculating the standard deviations among 100 data sets including 1000 samples that were selected at random. **a** Q_7 pair, **b** N_7 pair, **c** L_7 pair

basin of the aggregated state is wider and deeper than those for the other amino acids.

Figure 7 shows the distributions of the length of the monomer extension Λ_{mono} for REMD samples of Q_7, N_7, and L_7. For all samples during the simulations (Fig. 7a–c), there are two peaks at lengths of 11 and 20 Å. The peak at smaller values of Λ_{mono} corresponds to a shrunken structure, whereas the peak at larger values of Λ_{mono} corresponds to an extended structure. In the case of L_7, the peak height for the shorter structure is extremely high compared with other peptides. On the other hand, in the case of Q_7, the peak height for the shorter structure is low compared with that for the longer one. For D_{min} <6 Å (Fig. 7d–f), where the two monomers are considered to be dimerized, the peak height at longer Λ_{mono} increases for all models. In particular, there is no peak at the shorter Λ_{mono} in the results for Q_7 (Fig. 7d).

For each polypeptide, Fig. 8 shows the distribution of secondary structures for the dimerized samples in which the D_{min} value is less than 6 Å. The definitions of these structures are listed in Table 4. For all chain lengths, polyN dimers favor parallel β-sheet structures, whereas polyQ dimers favor antiparallel β-sheet structures. Notably, almost all of the Q_7 dimerized samples adopt antiparallel β-sheet structures. On the other hand, the L_7 dimers favor α-helix structures, as expected.

Figure 9 shows the number of intermolecular hydrogen bonds (n_{Hbond}^{inter}) for the dimerized samples of Q_7, N_7, and L_7. When the hydrogen atom (H) was located in between an oxygen atom (O) and another oxygen atom or a nitrogen atom (N), we assumed that the pair O–O or O–N was hydrogen bonded if the O–O or O–N distance was shorter than 3.3 Å and the O–H–O or O–H–N angle was between 120° and 180°. For the Q_7 dimer, the ratio of the n_{Hbond}^{inter} values for the main-chains accounted for 75 % of the hydrogen bonds. For the N_7 dimer, this ratio was about 50 %. These results indicate that the polyQ dimer is stabilized by hydrogen bonds between main-chain groups, whereas the polyN dimer has various combinations of intermolecular hydrogen bonds. For the L_7 dimer, the total value of n_{Hbond}^{inter} was smaller than the corresponding values for the other dimers.

MMD for Q_7 dimers

We performed MMD simulations for the Q_7 dimers starting from antiparallel or parallel initial structures to investigate which structure was the most stable. Figure 10a shows the proportions of MD runs in which the D_{min} values remained smaller than 6 Å. These ratios were calculated every 10 ns for each of the ten MD simulation runs. The simulations that started from the antiparallel structures exhibited an almost constant distance between the monomers in all ten MD runs during 30 ns of simulation time, whereas those which started from the parallel structures were found to dissociate in eight of the ten MD runs after 10 ns.

Figure 10b and c show representative structures after 10 ns MD runs in which the D_{min} values remained below 6 Å. In the snapshots starting from the antiparallel structure, the main chains of two monomers were aligned in straight lines, and the hydrogen bonds between the main chains were aligned in the same direction. In contrast, in the snapshots starting from the parallel structure, the chains were found to twist around each other.

Discussion

The sharp peaks in the distributions $p(D_{min})$ for pairs of polyQ and polyN (Fig. 4) indicate that the dimers of these peptides are bound by a short-range force. This short-range force must derive from electrostatic interactions involving hydrogen bonds between monomers. The energy barriers between the monomeric and dimeric states of these peptides, as shown in Fig. 6, may be attributed to the recombination

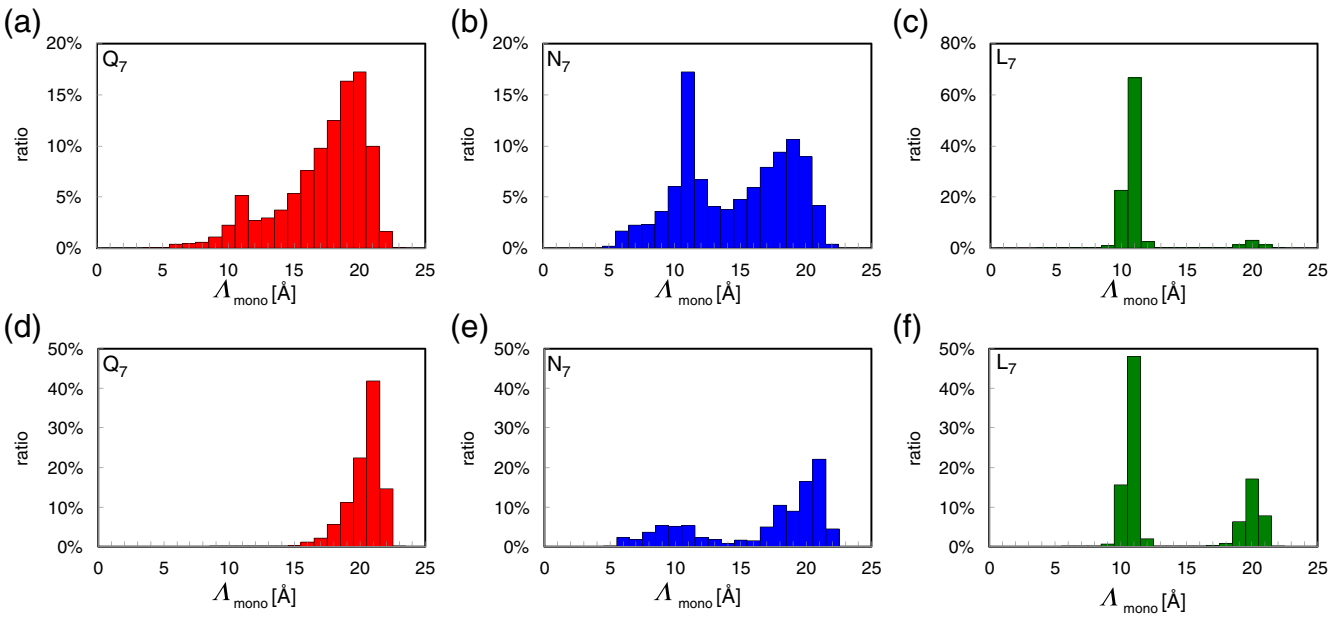

Fig. 7a–f Distributions of Λ_{mono} associated with the free-energy landscapes shown in Fig. 5. Here, Λ_{mono} is the length of the monomer extension, defined as the distance from the C_α atom of the C-terminus to that of the N-terminus. **a–c** Distributions of Λ_{mono} for all samples (D_{min} <40 Å). **d–f** Distributions of Λ_{mono} for the dimerized samples (D_{min} <6 Å)

of hydrogen bonds between peptide and water, or between two peptides.

However, the bonding patterns of the hydrogen bonds that formed between two Q_7 monomers were different from those of the hydrogen bonds that formed between two N_7 monomers. As seen in Fig. 9 for the Q_7 dimer, the number of hydrogen bonds ($n_{\mathrm{Hbond}}^{\mathrm{inter}}$) associated with the side chains is significantly smaller than the number of hydrogen bonds between the main chains. These results directly indicate that the Q_7 dimer is stabilized by hydrogen bonds between the main chains.

Wang and Voth [19] performed MD simulations using solvent-free multiscale coarse-graining models for various repeat lengths of polyQ peptides. They showed that the

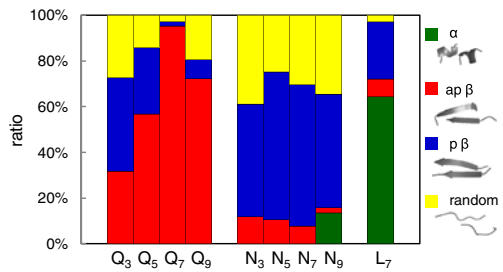

Fig. 8 Distribution of the secondary structures for polyQ dimers, polyN dimers, and the L_7 dimer. The dimerized samples comprise the snapshots in which the D_{min} value is less than 6 Å. These samples were taken from the replicas at 300 K for each model, and they are listed in Table 2. Definitions of these structures are provided in Table 4. *Green bar*: α-helix (*α*). *Red bar*: antiparallel β-sheet (*apβ*). *Blue bar*: parallel β-sheet (*pβ*). *Yellow bar*: random coil (*random*). Representative structures for each secondary structure that we defined are also shown

interactions associated with side chains are mainly responsible for peptide aggregation when the molecules are relatively far from each other (>5 Å), whereas the interaction between main chains is more important when the molecules are closer (i.e., around 3.5 Å). The roles of the hydrogen bonds obtained from our simulations are consistent with those in these previous simulations.

Figure 7a shows that monomeric polyQ peptides prefer to adopt extended structures, even when they are in the isolated state. On the other hand, Fig. 3 shows that the polyQ monomer also takes various structures, including a β-turn, whereas polyL or polyN exclusively adopt α-helical structures. These results suggest that polyQ peptides are primed in the isolated state to form a dimer with a β-sheet structure when another monomer approaches. To examine this hypothesis, we performed REMD simulations for a N_7 pair with values of Λ_{mono} constrained to be larger than 15 Å, which we refer to as N_7^{ext}. Figure 11 shows the distribution of D_{min} for the N_7^{ext} pair in comparison with those of N_7 and Q_7 pairs. The ratio of dimer formation increased dramatically when Λ_{mono} was constrained to be large.

Nakanishi and Kikuchi [34] investigated the mechanism for the aggregation of two proteins thermodynamically with the HP (hydrophobic/polar) lattice protein model. They suggested that the two peptide chains are primarily in their native state individually, but the native state becomes rapidly unstable when the distance between the two chains becomes shorter than the critical length at which these peptides make contact with each other. They also suggested that the lowest-energy state of the monomer should be

Table 4 Definitions of classes of secondary structure used for the dimerized samples in which the value of D_{min} is shorter than 6 Å. The program DPSS [33] was used to determine the secondary structure of each residue

Class	Definitions
α	Dimerized sample in which each monomer has more than one residue with properties of an α-helix
Parallel β-sheet (pβ)	Dimerized sample satisfies both of the following conditions:
	(1) Each monomer has more than one residue that has parallel β properties
	(2) The inner product of the vectors from the N-terminal to the C-terminal of each monomer is >0
Antiparallel β-sheet (apβ)	Dimerized sample satisfies both of the following conditions:
	(1) Each monomer has more than one residue that has antiparallel β properties
	(2) The inner product of the vectors from the N-terminal to the C-terminal of each monomer is <0.
Random	Dimerized sample that simultaneously satisfies the definitions of α and β, or satisfies neither of them

partially unfolded to become a dimer. Although their model peptides included only hydrophobic and polar residues, their folding scenario was similar to our results. Furthermore, Nagai et al. [35] showed that a β-sheet conformational transition of the expanded polyQ monomer precedes its assembly into β-sheet-rich amyloid-like fibrils. Our dimerization mechanism is consistent with their experimental results.

Figure 8 shows that the polyQ dimers strongly prefer the antiparallel β-sheet structure rather than the parallel β-sheet structure. These results are also supported by MMD calculations, as shown in Fig. 10.

Laghaei and Mousseau [20] performed REMD simulations with a coarse-grained force field to determine the structure of polyQ aggregates. In their study, the spontaneous formation of antiparallel β-sheets as well as triangular and circular β-helical structures for polyQ peptides with a length of 40 residues was observed. Since the lengths of the peptides in our calculation were short, it is difficult to determine whether the aggregate form is a sheet or helix structure. However, our results indicate that polyQ peptides strongly favor the antiparallel β-sheet structure over the parallel structure. This result is consistent with the above-mentioned coarse-grained simulations.

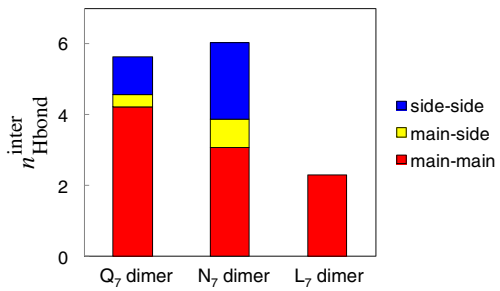

Fig. 9 Number of intermolecular hydrogen bonds (n_{Hbond}^{inter}) for the Q_7, N_7, and L_7 dimers at 300 K. When the hydrogen atom (H) was located between an oxygen atom (O) and another oxygen atom or a nitrogen atom (N), we assumed that the O–O or O–N pair was hydrogen bonded if the O–O or O–N distance was shorter than 3.3 Å and the O–H–O or O–H–N angle was between 120 and 180°. *Red bar*: hydrogen bonds between the main chains. *Yellow bar*: those between the main chains and the side chains. *Blue bar*: those between the side chains

The fact that polyQ peptides prefer an antiparallel β-sheet structure indicates that the monomeric polyQ peptide can form antiparallel β-sheet structures by itself. We showed in our previous study [18] that an oligomer consisting of polyQ monomers, which adopt β-turn structures, is stable compared to the monomeric β-turn structure. These findings provide us with a clue to unlocking the mechanism of nucleation.

A long polyQ peptide can form stable structures by making several β-turn segments within the chain. Since

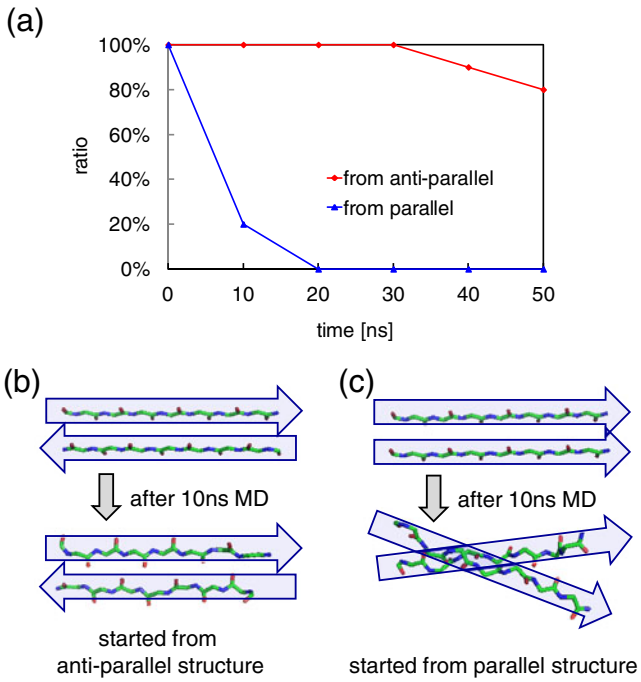

Fig. 10 **a** Proportions (i.e., ratios) of ten MD runs in which D_{min} values <6 Å during the simulations. *Red line*: started from the antiparallel structure. *Blue line*: started from the parallel structure. **b**, **c** Initial and final structures of the Q_7 dimer for 10 ns MD simulations in which D_{min} was <6 Å. **b** Started from the antiparallel structure. **c** Started from the parallel structure. Carbon atoms are *green*, oxygen atoms are *red*, and nitrogen atoms are *blue*. Hydrogen atoms and side chains are omitted for clarity. *Arrows* along the Q_7 monomers indicate the directions of the peptide chains (from N-termini to C-termini)

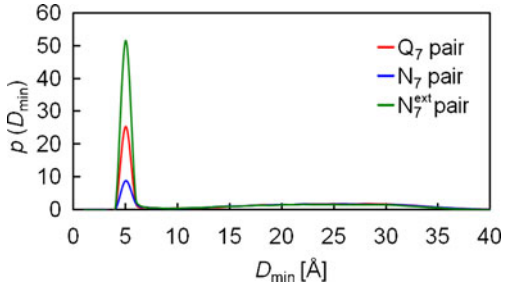

Fig. 11 Normalized distribution function $p(D_{min})$ as a function of D_{min} for N_7^{ext}, N_7, and Q_7 pairs at 300 K

such a structure is relatively stable, and thus gives rise to β-strands which have high binding affinity, this peptide conformation can recruit the next polyQ peptide. This long polyQ peptide can then act as a seed to facilitate aggregate formation. We present a scheme of the nucleation mechanism for polyQ peptides based on the results of this study in Fig. 12.

Interestingly, we found some evidence that seven is a special number for the polyQ peptides in this study. Figure 3b shows that the β-strand always turns at the eighth glutamine residue, regardless of the chain length. Although it is difficult to compare the peak heights among different models because of differences in the number of atoms, the degree of dimerization of Q_7 in Fig. 4a is significantly higher than it is for the other lengths. Figure 8 shows that the ratio of the antiparallel β-sheet structure of the Q_7 dimer is higher than the ratios for other chain lengths.

Thakur and Wetzel [36] proposed that the ideal folding motif for polyQ aggregates consists of alternating the extended chains of seven residues. Lakhani et al. [21] showed that the lengths of β-strands within a polyQ peptide range from six to nine residues using replica exchange discrete molecular dynamics simulations. These results indicate that a seven-glutamine repeat is a fundamental element of the β-turn structure. Our simulation results lend further support to these findings. However, we still cannot explain why polyQ prefers the number seven. If we could determine the number of residues that are needed to form stable β-strands and connect them, we could estimate the minimum length that a polyQ peptide needs to be in order to act as an aggregation nucleus. This is a problem to solve in the future.

Conclusions

The research objectives of this study were to estimate the energy barrier between the monomeric and dimeric states of polyQ peptides, to clarify the reasons for the substantial ability of polyQ peptides to aggregate, and to elucidate the dynamical processes involved in the nucleation and aggregation of polyQ peptides. To answer these questions, we performed REMD and MMD simulations to explore the conformational ensemble of polyQ peptides over a wide range of energy space. Furthermore, to determine the characteristics of the polyQ peptide, we compared several properties of polyQ with those of different peptides composed of other amino acids. Based on our analyses of these simulations, we have obtained several new findings, as follows.

On the free-energy landscape of polyQ peptides, there is a low-energy barrier (~2 kcal/mol) between the monomeric and dimeric states, and this barrier height is low enough that it can be overcome at physiological temperatures.

When glutamine residues approach each other and reach a distance where hydrogen bonds can form between residues, polyQ peptides extend themselves and readily adopt antiparallel β-sheet structures. We also found that the antiparallel β-sheet structure is stable because the main chains of polyQ peptides can adopt an extended straight conformation by forming parallel hydrogen bonds.

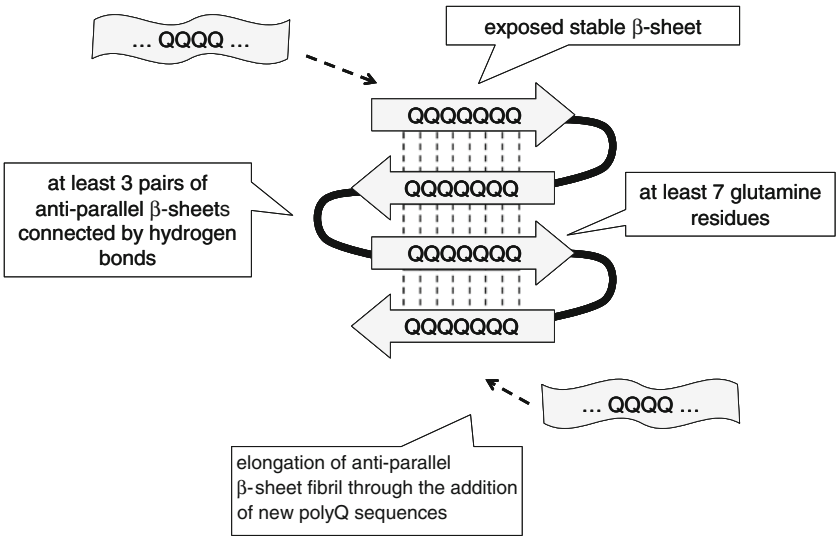

Fig. 12 Schematic of the proposed nucleation and elongation process of the polyQ peptide

A long polyQ peptide can form stable structures by introducing several β-turn segments into the chain. If we could determine the number of residues needed to form stable β-strands and link them together, we could predict the minimum length that a polyQ peptide needs to be for it to act as an aggregation nucleus.

The polyQ peptide is unique because it consists only of glutamines. However, this peptide shares some characteristics with the peptides associated with amyloid formation [37]. Consequently, our study yields an explanation of the probable mechanism for the dimerization of polyQ peptides, which may also provide insight into the general scenario of the amyloid fibril formation process.

Acknowledgments We thank Dr. Takatoshi Fujita for useful comments and discussions. The numerical calculations were partially carried out on a PC-cluster system in the Cybermedia Center at Osaka University.

References

1. The Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72:971–983. doi:10.1016/0092-8674(93)90585-E
2. Rubinsztein DC, Leggo J, Coles R, Almqvist E, Biancalana V, Cassiman JJ, Chotai K, Connarty M, Crauford D, Curtis A, Curtis D, Davidson MJ, Differ AM, Dode C, Dodge A, Frontali M, Ranen NG, Stine OC, Sherr M, Abbott MH, Franz ML, Graham CA, Harper PS, Hedreen JC, Jackson A, Kaplan JC, Losekoot M, MacMillan JC, Morrison P, Trottier Y, Novelletto A, Simpson SA, Theilmann J, Whittaker JL, Folstein SE, Ross CA, Hayden MR (1996) Phenotypic characterization of individuals with 30–40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36–39 repeats. Am J Hum Genet 59:16–22
3. Chen S, Ferrone FA, Wetzel R (2002) Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. Proc Natl Acad Sci USA 99:11884–11889. doi:10.1073/pnas.182276099
4. Cooper JK, Schilling G, Peters MF, Herring WJ, Sharp AH, Kaminsky Z, Masone J, Khan FA, Delanoy M, Borchelt DR, Dawson VL, Dawson TM, Ross CA (1998) Truncated N-terminal fragments of huntingtin with expanded glutamine repeats form nuclear and cytoplasmic aggregates in cell culture. Hum Mol Genet 7:783–790. doi:10.1093/hmg/7.5.783
5. Chen S, Berthelier V, Hamilton JB, O'Nuallain B, Wetzel R (2002) Amyloid-like features of polyglutamine aggregates and their assembly kinetics. Biochemistry 41:7391–7399. doi:10.1021/bi011772q
6. Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, Hasenbank R, Bates GP, Davies SW, Lehrach H, Wanker EE (1997) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell 90:549–558. doi:10.1016/S0092-8674(00)80514-0
7. Hollenbach B, Scherzinger E, Schweiger K, Lurz R, Lehrach H, Wanker EE (1999) Aggregation of truncated GST-HD exon 1 fusion proteins containing normal range and expanded glutamine repeats. Philos Trans R Soc Lond B Biol Sci 354:991–994
8. Georgalis Y, Starikov EB, Hollenbach B, Lurz R, Scherzinger E, Saenger W, Lehrach H, Wanker EE (1998) Huntingtin aggregation monitored by dynamic light scattering. Proc Natl Acad Sci USA 95:6118–6121
9. Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, Bates GP, Lehrach H, Wanker EE (1999) Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. Proc Natl Acad Sci USA 96:4604–4609. doi:10.1073/pnas.96.8.4604
10. Perutz MF, Johnson T, Suzuki M, Finch JT (1994) Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proc Natl Acad Sci USA 91:5355–5358
11. Sharma D, Sharma S, Pasha S, Brahmachari SK (1999) Peptide models for inherited neurodegenerative disorders: conformation and aggregation properties of long polyglutamine peptides with and without interruptions. FEBS Lett 456:181–185. doi:10.1016/S0014-5793(99)00933-3
12. Tanaka M, Morishima I, Akagi T, Hashikawa T, Nukina N (2001) Intra- and intermolecular beta-pleated sheet formation in glutamine-repeat inserted myoglobin as a model for polyglutamine diseases. J Biol Chem 276:45470–45475. doi:10.1074/jbc.M107502200
13. Bevivino AE, Loll PJ (2001) An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel beta-fibrils. Proc Natl Acad Sci USA 98:11955–11960. doi:10.1073/pnas.211305198
14. Sharma D, Shinchuk LM, Inouye H, Wetzel R, Kirschner DA (2005) Polyglutamine homopolymers having 8–45 residues form slablike beta-crystallite assemblies. Proteins 61:398–411. doi:10.1002/prot.20602
15. Vitalis A, Wang X, Pappu RV (2008) Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. J Mol Biol 384:279–297. doi:10.1016/j.jmb.2008.09.026
16. Crick SL, Jayaraman M, Frieden C, Wetzel R, Pappu RV (2006) Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. Proc Natl Acad Sci USA 103:16764–16769. doi:10.1073/pnas.0608175103
17. Masino L, Kelly G, Leonard K, Trottier Y, Pastore A (2002) Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. FEBS Lett 513:267–272. doi:10.1016/S0014-5793(02)02335-9
18. Nakano M, Watanabe H, Rothstein SM, Tanaka S (2010) Comparative characterization of short monomeric polyglutamine peptides by replica exchange molecular dynamics simulation. J Phys Chem B 114:7056–7061. doi:10.1021/jp9122024
19. Wang Y, Voth GA (2010) Molecular dynamics simulations of polyglutamine aggregation using solvent-free multiscale coarse-grained models. J Phys Chem B 114:8735–8743. doi:10.1021/jp1007768
20. Laghaei R, Mousseau N (2010) Spontaneous formation of polyglutamine nanotubes with molecular dynamics simulations. J Chem Phys 132:165102. doi:10.1063/1.3383244
21. Lakhani VV, Ding F, Dokholyan NV (2010) Polyglutamine induced misfolding of huntingtin exon1 is modulated by the flanking sequences. PLoS Comput Biol 6:e1000772. doi:10.1371/journal.pcbi.1000772
22. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 314:141–151. doi:10.1016/S0009-2614(99)01123-9
23. Caves LS, Evanseck JD, Karplus M (1998) Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin. Protein Sci 7:649–666
24. Straub JE, Rashkin AB, Thirumalai D (1994) Dynamics in rugged energy landscapes with applications to the S-peptide and ribonuclease A. J Am Chem Soc 116:2049–2063. doi:10.1021/ja00084a051

25. Elofsson A, Nilsson L (1993) How consistent are molecular dynamics simulations? Comparing structure and dynamics in reduced and oxidized *Escherichia coli* thioredoxin. J Mol Biol 233:766–780. doi:10.1006/jmbi.1993.1551
26. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, Merz KM, Wang B, Hayik S, Roitberg A, Seabra G, Kolossvary I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V, Kollman PA (2008) AMBER 10. University of California San Francisco, San Francisco
27. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins 55:383–394. doi:10.1002/prot.20033
28. Zhou R (2003) Free energy landscape of protein folding in water: explicit vs. implicit solvent. Proteins 53:148–161. doi:10.1002/prot.10483
29. Shell MS, Ritterson R, Dill KA (2008) A test on peptide stability of AMBER force fields with implicit solvation. J Phys Chem B 112:6878–6886. doi:10.1021/jp800282x
30. Gnanakaran S, Nymeyer H, Portman J, Sanbonmatsu KY, García AE (2003) Peptide folding simulations. Curr Opin Struct Biol 13:168–174. doi:10.1016/S0959-440X(03)00040-X
31. Schneider T, Stoll E (1978) Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. Phys Rev B 17:1302–1322. doi:10.1103/PhysRevB.17.1302
32. Pastor RW, Brooks BR, Szabo A (1988) An analysis of the accuracy of Langevin and molecular dynamics algorithms. Mol Phys 65:1409–1419. doi:10.1080/00268978800101881
33. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637. doi:10.1002/bip.360221211
34. Nakanishi K, Kikuchi M (2006) Thermodynamics of aggregation of two proteins. JPSJ 75:64803–64806. doi:10.1143/JPSJ.75.064803
35. Nagai Y, Inui T, Popiel HA, Fujikake N, Hasegawa K, Urade Y, Goto Y, Naiki H, Toda T (2007) A toxic monomeric conformer of the polyglutamine protein. Nat Struct Mol Biol 14:332–340. doi:10.1038/nsmb1215
36. Thakur AK, Wetzel R (2002) Mutational analysis of the structural organization of polyglutamine aggregates. Proc Natl Acad Sci USA 99:17014–17019. doi:10.1073/pnas.252523899
37. Nakano M (2010) Doctoral thesis. Kobe University, Kobe Japan
38. DeLano Scientific LLC (2006) PyMOL viewer v0.99. DeLano Scientific LLC, Palo Alto