J. Biochem. **136**, 583–594 (2004)
DOI: 10.1093/jb/mvh165

General Dynamic Properties of $A\beta_{12-36}$ Amyloid Peptide Involved in Alzheimer's Disease from Unfolding Simulation

Shinya Suzuki[1], Oxana V. Galzitskaya[2], Daisuke Mitomo[1] and Junichi Higo[1,3,*]

[1]Laboratory of Bioinformatics, School of Life Science, Tokyo University of Pharmacy and Life Science, 1432-1 Horinouchi, Hachioji, Tokyo, 192-0392; [2]Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290, Russia; and [3]BIRD, JST (Japan Science and Technology Corporation)

Received June 19, 2004; accepted August 11, 2004

To study the folding/unfolding properties of a β-amyloid peptide $A\beta_{12-36}$ of Alzheimer's disease, five molecular dynamics simulations of $A\beta_{12-36}$ in explicit water were done at 450 K starting from a structure that is stable in trifluoroethanol/water at room temperature with two α-helices. Due to high temperature, the initial helical structure unfolded during the simulation. The observed aspects of the unfolding were as follows. 1) One helix (helix 1) had a longer life than the other (helix 2), which correlates well with the theoretically computed Φ values. 2) Temporal prolongation of helix 1 was found before unfolding. 3) Hydrophobic cores formed frequently with rearrangement of amino-acid residues in the hydrophobic cores. The formation and rearrangement of the hydrophobic cores may be a general aspect of this peptide in the unfolded state, and the structural changes accompanied by the hydrophobic-core rearrangement may lead the peptide to the most stable structure. 4) Concerted motions (collective modes) appeared to unfold helix 1. The collective modes were similar with those observed in another simulation at 300 K. The analysis implies that the conformation moves according to the collective modes when the peptide is in the initial stage of protein unfolding and in the final stage of protein folding.

Key words: collective mode, hydrophobic core, molecular dynamics, principal component analysis, transition state.

Abbreviations: Aβ,β-amyloid peptide; H-bond, hydrogen bond; MD, molecular dynamics; PCA, principal component analysis.

One of the most important subjects in protein biochemistry is a rational understanding or prediction of protein functioning. Since protein function is a property assigned to the native structure, studies on the native structure are more informative than those on protein folding, although protein folding has been a fundamentally important issue in biophysics. However, amyloidosis proposes a problem deeply related to the folding mechanism of a protein/peptide (*1*).

It is believed that fibril formation is a key event in amyloidosis, in which proteins/peptides are accumulated in a specific β form. In this sense, amyloidosis is an issue of molecular interactions. However, fibril formation is also a process in which the protein conformation changes from the native structure to the β-structure. Experiments (*2–5*) have shown that the fibril formation depends on the solution conditions, such as pH, and have suggested that some denaturation processes of protein structure facilitate fibril formation. A simulation study has shown that the foldability of a polypeptide chain into different structures may be important at the initial stage of the fibril formation (*6*). Exhaustive conformational sampling has demonstrated that short peptides generally tend to fold into different secondary structures in water (*7–9*). A bioinformatic approach showed that some amyloid-formable proteins involve a helical region that was predicted to be a β-strand by secondary-structure prediction methods (*10*). It has been suggested that a polypeptide chain may generally have an ability to fold into a non-native structure depending on the solution conditions (*11*). These studies suggest that the single-chain folding mechanism is also important for understanding the initial state of fibril formation.

The structural heterogeneity of the β-amyloid peptide, Aβ, of Alzheimer's disease has been observed by NMR study. In a membrane-mimicking solution (a mixture of trifluoroethanol and water), the structure of Aβ has helical elements (*12–17*). We call this conformation a "helix-involving structure" in this study. Trifluoroethanol induces secondary-structure formation in polypeptides (*18*), and weakens hydrophobic interactions in the protein tertiary structure (*19*). The sequence of Aβ has two highly hydrophobic regions, each of which folds into an α-helix in a mixture of trifluoroethanol/water (*17*). On the other hand, Aβ has a "collapsed-coil structure" in water (*20*). Solid-state NMR measurements have shown that Aβ can accumulate in a parallel β-sheet (*21*, *22*). This structural heterogeneity may be of crucial importance for amiloidosis, because the structural transition may play an essential role in fibril formation (*6*).

It is difficult to experimentally measure the structural changes of Aβ in fibril formation at an atomic resolution. A computational method, molecular dynamics (MD) simulation, may provide useful information on the structural

*To whom correspondence should be addressed. Tel: +81-462-76-5498, Fax: +81-462-76-5351, E-mail: higo@ls.toyaku.ac.jp

changes, because by this method, it is possible to trace protein conformational changes at atomic resolution. However, the conventional MD simulation is limited in time, which is shorter than several tens of nanoseconds in most cases. This means that the observation of fibril formation is beyond of the limit of the MD simulation. On the other hand, the inverse process (*i.e.*, dissociation) of fibril formation may be reproduced by MD simulation, because the process does not take a long computational time when the simulation conditions are appropriately or carefully set to reproduce the dissociation. In fact, dissociation processes of fibrils have been studied by MD simulation: some computer-modeled Aβ oligomers were tested by MD simulations, and the stable structure of the oligomers during the MD simulation was proposed to be a possible fibril structure (*23*).

A similar situation exists in the study of protein folding: although it is still difficult to observe the folding process in MD simulation, unfolding can be reproduced at a high temperature. Note that the higher the temperature, the faster the unfolding of the protein native structure. High-temperature MD simulation is generally called unfolding MD simulation. This simulation method provides a picture of the transition-state ensemble of proteins, and the simulation results correlate well with the experimental Φ values (*24–26*). Amino-acid residues with high Φ values form a folding nucleus at the transition state in refolding/unfolding (*27*). The simulation and experiment are combined to study the refolding/unfolding pathways of proteins (*28*, *29*).

We did high-temperature MD simulation to trace the unfolding of Aβ in water. The tertiary structure from which the simulation starts (i.e., the starting point of simulation) is called the "initial structure." The choice of the initial structure is critically important because the unfolding MD simulation produces unfolding of the initial structure. In this study, we chose the helix-involving structure as the initial structure. Since the most stable structure of Aβ in water is not the helix-involving structure but the collapsed-coil structure (*12*, *20*), the helix-involving structure should unfold spontaneously in water during MD simulation. However, the helix-involving structure was also stable during a 1-ns MD simulation in water at 300 K (see "RESULTS"). A similar apparent stability of the initial structure, which may be unstable under the given simulation conditions, is often observed in a room-temperature MD simulation, because unfolding at room temperature occurs in a time scale too long to be reproduced by the MD simulation. We were interested in the helix-involving structure because the simulation study (*6*) inferred a possibility that the helical conformation may be intermediate before fibril formation. Besides, a helix is a commonly found structural element in proteins. Thus, if there is a typical conformational motion to break the helix, the unfolding of the helix-involving structure may be useful for studying the unfolding of other proteins. Therefore, we chose the helix-involving structure as the initial structure, and observed the general properties of amyloid peptide $A\beta_{12-36}$ of Alzheimer's disease in unfolding simulations. 1) One helix (helix 1) had a longer life than the other (helix 2), which correlates well with the theoretically computed Φ values. 2) Temporal prolongation of helix 1 was found before unfolding. 3) Hydrophobic cores frequently formed with rearrangement of amino-acid residues in the hydrophobic cores. 4) Collective motions (collective modes) appeared to unfold helix 1.

MATERIALS AND METHODS

The peptide Aβ originally consists of 40 amino-acid residues: [Asp-Ala-Glu-Phe-Arg-His-Asp-Ser-Gly-Tyr-Glu-Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-Gly-Ser-Asn-Lys-Gly-Ala-Ile-Ile-Gly-Leu-Met-Val-Gly-Gly-Val-Val], which is usually expressed as $A\beta_{1-40}$ to clearly specify the sequence. As explained in INTRODUCTION, Aβ exhibits structural heterogeneity (*12*, *20–22*), and a single chain can have two tertiary structures in solution: the collapsed-coil structure in water (*20*) and the helix-involving structure in 40 % (by volume) trifluoroethanol/water (*12*). We designate the former structure as C_{wat} and the latter as C_{trf}. We chose C_{trf} for the initial structure for the unfolding MD simulation: C_{trf} = MODEL 1 of 20 deposited NMR conformations in pdb (pdb entry = 1AML). We call the two helices in C_{trf} "helix 1" and "helix 2": helix 1 is the longer helix comprising residues 12 to 22 of $A\beta_{1-40}$, and helix 2 in the shorter helix comprising residues 31 to 36. Residues 23 to 30, linking the two helices, form a coil region. The N-terminal eleven residues and the C-terminal four residues do not converge well in the 20 deposited conformations.

Before MD simulation, we removed the non-converged N- and C-terminal regions of C_{trf}, to reduce the computational task. Thus, the computed sequence was $A\beta_{12-36}$ = [Ace-^{12}Val-^{13}His-^{14}His-^{15}Gln-^{16}Lys-^{17}Leu-^{18}Val-^{19}Phe-^{20}Phe-^{21}Ala-^{22}Glu-^{23}Asp-^{24}Val-^{25}Gly-^{26}Ser-^{27}Asn-^{28}Lys-^{29}Gly-^{30}Ala-^{31}Ile-^{32}Ile-^{33}Gly-^{34}Leu-^{35}Met-^{36}Val-Nme], where Ace and Nme are the N-terminal acetyl and C-terminal N-methyl groups, respectively, added to block the charges at the peptide termini. Nearly the same sequence (*i.e.*, $A\beta_{10-35}$) was studied for the experimental structural determination (*20*, *21*).

An MD simulation generally provides a number of snapshots, in which the positions of atoms in the polypeptide and water molecules are recorded. The ensemble of snapshots is called an MD trajectory, from which the changes in atomic positions (*i.e.*, conformational motions of molecules) starting from the initial structure can be monitored. In the current study, we carried out high-temperature unfolding MD simulations of C_{trf}. Thus, the MD trajectory produced the unfolding processes of C_{trf}.

$A\beta_{12-36}$ should be surrounded by water during the MD simulation. Then, the following computational procedure was used. First, the initial structure (*i.e.*, C_{trf}) was immersed in a sphere (radius 23.5 Å) consisting of water molecules, with the geometrical center of C_{trf} set at the center of the sphere. This means that C_{trf} was located at the center of the water sphere. There were 1,636 water molecules and 397 peptide atoms (total 5,305 atoms). In addition, the momentum and angular momentum of $A\beta_{12-36}$ were constrained at zero during the simulation. This means that the position and orientation of the polypeptide did not change in the water sphere during the simulation. Second, a harmonic potential was applied to the water molecules only when the molecules flew to the outside of the water sphere. This means that the

water molecules are confined in the sphere during the simulation. Third, the harmonic potential was also applied on peptide atoms only when they flew to the outside of a smaller sphere (radius = 20.0 Å), for which the center was set at the center of the water sphere. As a result, all peptide atoms were always wrapped by water during the simulation. The computational system (*i.e.*, the atomic positions in the polypeptide and the water molecules) generated here was used for the initial structure of an MD simulation at room temperature, which was done prior to the high-temperature unfolding MD simulation, and the room-temperature simulation was used to prepare the initial structures of the high-temperature unfolding simulations (see below).

All of the MD simulations were done with the PRESTO computer program (*30*) ver. 3. In the MD simulation, the atom-atom interactions are calculated with mathematical functions called force-field parameters. We used the AMBER parm 96 (*31*) force-field parameters for the peptide atoms and the TIP3P potential functions (*32*) for the water molecules. The choice of force field is essential when the protein folding/unfolding is studied, and it has been proved (*7*, *33*) that AMBER parm 96 is more appropriate than parm 94 (*34*). If conformational fluctuations around a well-defined native structure are studied, the choice of force fields is less essential, because the overall conformational fluctuations are roughly determined by the backbone folded pattern (*35*), which is not the case in the present study. The other computational details adopted are as follows: the cell-multipole expansion method (*36*) to speed up the computation of long-range electrostatic interactions; the constant-temperature method (*37*) to control the temperature; the SHAKE algorithm (*38*) to constrain covalent bonds with hydrogen atoms; 1.0-fs time step to solve the equation of motion.

Before the high-temperature unfolding simulation, an MD simulation was done performed at room temperature (300K) for 1.0 ns, using the initial structure described above with C_{trf} immersed in the water sphere. We call this simulation "300 K simulation." Since C_{trf} was stable during the 300 K simulation, as shown below, the trajectory (300 K trajectory) provided conformations fluctuating only around C_{trf}. Thus, the 300 K trajectory is a good control to be used for comparison of trajectories obtained from the high-temperature unfolding simulation.

The unfolding of a small protein generally involves multiple pathways in the conformational space (*39*). This means that multiple simulations are required to study the unfolding, where each simulation possibly produces a different unfolding pathway. To produce a variety of unfolding processes, we prepared five different initial structures for the unfolding simulation. Note that the five initial structures should be close to C_{trf}, because the current study focuses on the unfolding of the helix-involving structure. We chose five conformations from the 300 K trajectory at 0.2, 0.4, 0.6, 0.8, and 1.0 ns, and started five unfolding MD simulations at 450 K starting from these conformations. The five unfolding simulations were called ST2 (the initial structure is the conformation at 0.2 ns of the 300K trajectory), ST4 (0.4 ns), ST6 (0.6 ns), ST8 (0.8 ns), and ST10 (1.0 ns). A snapshot was saved once every 1,000 steps (=1.0 ps).

Before performing the unfolding MD simulations, we had confirmed that both structures C_{wat} and C_{trf} were unfolded at 450 K by making preliminary MD simulations. This means that a temperature of 450 K initiates the unfolding processes of $A\beta_{12-36}$.

The α-helices were stabilized by intra-main chain hydrogen bonds (H-bonds) $^{i}O\text{-}^{i+4}H$, which represent the H-bond between the carbonyl oxygen of the *i*'th residue and the amide hydrogen of the (*i+4*)'th residue, where i = 12,…, 18 (for helix 1) and i = 31,32 (for helix 2). Helix 1 in C_{trf} involves seven H-bonds. Here, $^{12}O\text{-}^{16}H$ was designated as HB_1 and the distance as r_1; $^{13}O\text{-}^{17}H$ as HB_2 and the distance as r_2; $^{14}O\text{-}^{18}H$ as HB_3 and the distance as r_3; $^{15}O\text{-}^{19}H$ as HB_4 and the distance as r_4; $^{16}O\text{-}^{20}H$ as HB_5 and the distance as r_5; $^{17}O\text{-}^{21}H$ as HB_6 and the distance as r_6; and $^{18}O\text{-}^{22}H$ as HB_7 and r_7. In helix 2, there are two H-bonds: $^{31}O\text{-}^{35}H$ was designated as HB_8 and the distance as r_8; and $^{32}O\text{-}^{36}H$ as HB_9 and the distance as r_9. Note that the distances [r_1, …, r_9] can be calculated even when the H-bonds are broken. We designated the region from residue 12 to 22 as "helix-1 region" and that from residue 31 to 36 as "helix-2 region." These designations were used even when the helices were broken.

We continued each unfolding MD simulation until all of the H-bonds were broken. Thus, the simulation lengths differed among the five simulations. As shown in RESULTS, the conformation of $A\beta_{12-36}$ became random in each unfolding simulation.

Atomic positions determine the polypeptide conformation. If there are N atoms in the polypeptide (N = 397 for the current study), then the number of degrees of freedom to describe the polypeptide conformation is 3N. This means that a polypeptide conformation is assigned to a position in the 3N dimensional space. On the other hand, the multi-dimensionality makes the conformational analysis difficult, because one can image only one-, two-, or three-dimensional (3D) space. Thus, an effective method to treat the multi-dimensional space is crucially important. The principal component analysis (PCA) has been widely used for protein conformational and dynamic analysis (*40–47*), because the method can pick up a few essential dimensions (*i.e.*, coordinate axes) from the 3N dimensional space. When three axes are picked up, the axes can construct a 3D space, and the polypeptide conformations can be projected on the 3D space. This means that the conformational distribution is visible in the 3D space.

In this study, we used PCA to analyze the unfolding of helix 1 in the unfolding simulation. The conformational motions (*i.e.*, snapshots) of $A\beta_{12-36}$ are recorded in the MD trajectory, and the unfolding of helix 1 can be monitored by a set of seven lengths [r_1, …, r_7], which is expressed by a seven-dimensional (7D) vector as $\boldsymbol{R}$ = [r_1, …, r_7]. The PCA starts by calculating a matrix (a valence-covariant matrix): $C_{ij} = \langle q_i\, q_j \rangle - \langle q_i \rangle \langle q_j \rangle$, where C_{ij} is the *i,j*'th element of the matrix, and <…> represents the trajectory average (i.e., the average over all snapshots in the MD trajectory). The quantities q_i and q_j are the *i*'th and *j*'th elements of an arbitrary vector derived from the polypeptide conformation. Thus, the selection of variables [q_1, q_2, q_3, …] taken into the matrix is essentially important to elucidate structural properties of the conformational motions of the polypeptide. We used seven lengths (*i.e.*, $q_1 \rightarrow r_1$, …, $q_7 \rightarrow r_7$) for calculating the matrix as follows:

$$C_{ij} = \langle r_i\, r_j\rangle - \langle r_i\rangle \langle r_j\rangle, \quad (1)$$

where $\langle r_i\, r_j\rangle$ is the trajectory average of $r_i\, r_j$, and $\langle r_i\rangle$ is that of r_i (*i.e.*, the average length of r_i). The matrix size is 7×7, because both i and j range from 1 to 7 in C_{ij}. By diagonalizing the matrix, eigen-vectors, $\boldsymbol{v}_1$, …, $\boldsymbol{v}_7$, and eigen-values, λ_1, …, λ_7, were obtained. The eigen-vectors satisfy the equation: $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \delta_{ij}$, which means that the eigen-vectors can be regarded as coordinate axes (i.e., normalized and orthogonal coordinate axes) to construct the 7D space. The $\boldsymbol{v}_i$ is expressed with the elements as $[v_{1i}, \ldots, v_{7i}]$, where v_{ki} is the k'th element of $\boldsymbol{v}_i$. We designated the 7D space constructed by $\boldsymbol{v}_1$, …, $\boldsymbol{v}_7$ as $\Gamma_{7\mathrm{HB}}$.

Now, we explain how to express or assign the sampled conformations into $\Gamma_{7\mathrm{HB}}$. Given the conformation, $\boldsymbol{R}'$, of the helix-1 region: $\boldsymbol{R}' = [r'_1, \ldots, r'_7]$. Then, $\boldsymbol{R}'$ can be described by $\boldsymbol{v}_i$.

$$\boldsymbol{R}' = \Sigma_i\, c_i \boldsymbol{v}_i + \langle \boldsymbol{R}\rangle, \quad (2)$$

where

$$c_i = (\boldsymbol{R}' - \langle \boldsymbol{R}\rangle)\cdot \boldsymbol{v}_i = \Sigma_m\, (r'_m - \langle r_m\rangle) v_{mi}. \quad (3)$$

Equation 2 is a well-used expression (*i.e.*, linear combination) to represent a vector by unit-vectors: for example, the position $[x, y, z]$ in a real space is expressed as $x\boldsymbol{e}_x + y\boldsymbol{e}y_x + z\boldsymbol{e}_z$, where $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, and $\boldsymbol{e}_z$ are unit-vectors corresponding to the x-, y-, and z-axes, respectively, and then x, y, and z are the projections of the position onto the axes. In Eq. 2, the unit-vectors are eigen-vectors $\boldsymbol{v}_1$, …, $\boldsymbol{v}_7$, and c_i is the projection of $\boldsymbol{R}' - \langle \boldsymbol{R}\rangle$ on $\boldsymbol{v}_i$. The origin of $\Gamma_{7\mathrm{HB}}$ is set on the mean $\langle \boldsymbol{R}\rangle$. This means that the seven lengths are measured from the mean lengths $\langle \boldsymbol{R}\rangle = [\langle r_1\rangle, \ldots, \langle r_7\rangle]$.

As mentioned above, the conformation of the helix-1 region is described by the set $[r_1, \ldots, r_7]$. The unfolding MD simulations generated a variety of conformations of the helix-1 region. If the lengths are close to the H-bond length, the conformation is helical, and if not, the helix is broken. The conformational distribution for the helix-1 region in $\Gamma_{7\mathrm{HB}}$ is obtained as follows: when projecting a conformation in the MD trajectory with Eq. 3, a set of coefficients $[c_1, \ldots, c_7]$ is obtained. Then, from all of the conformations, an ensemble of coefficients is obtained: $[c_1{}^1, \ldots, c_7{}^1]$, $[c_1{}^2, \ldots, c_7{}^2]$, …, $[c_1{}^S, \ldots, c_7{}^S]$, where the superscripts represent the ordinal numbers assigned to the snapshots (*i.e.*, c_i^k is the i'th coefficient in the k'th snapshot in the trajectory), and S is the number assigned to the last snapshot in the trajectory.

Since the conformations of the helix-1 region were defined in the 7D space, the conformational distribution is invisible. To view the distribution, we shose the first, second, and third elements from the 7D representation: $[c_1{}^1, c_2{}^1, c_3{}^1]$, $[c_1{}^2, c_2{}^2, c_3{}^2]$, …, $[c_1{}^S, c_2{}^S, c_3{}^S]$. There are other ways to select the three elements rather than from the first to the third ones. However, in PCA, the standard deviation of the distribution along $\boldsymbol{v}_i$ is given by λ_i, which means that the larger the eigen-value, the wider the distribution along the eigen-vector. Here, we arranged the eigen-values in descending order. Thus, the selection of the first to third elements is plausible to view the details of the distribution.

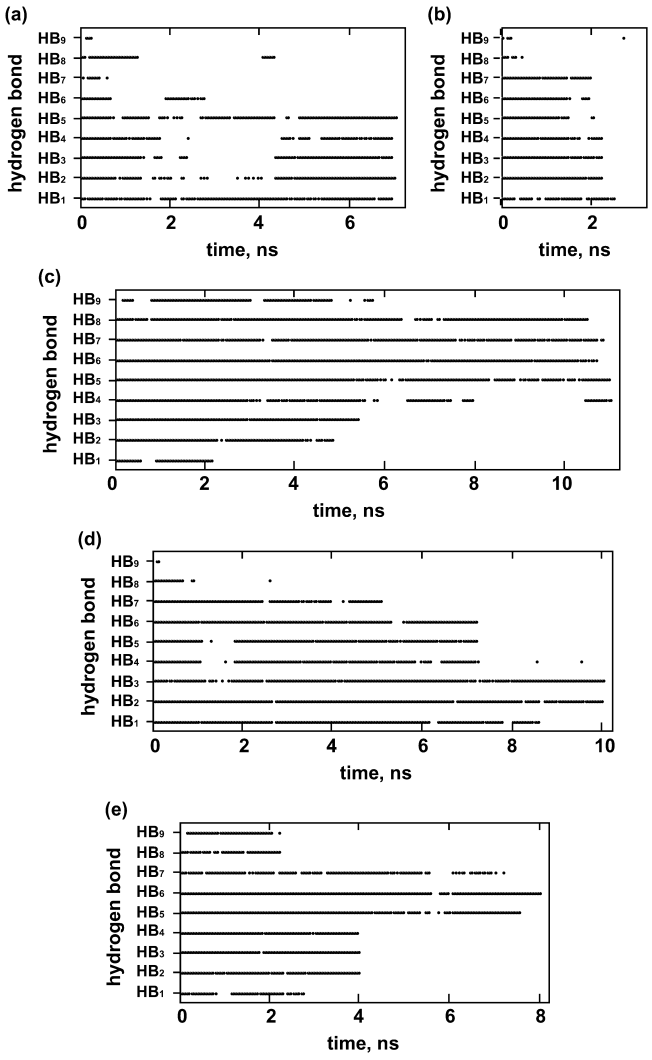


Fig. 1. **Breaking of HB_1, …, HB_9.** If r_i is smaller than 2.5 Å at a given time, HB_i is judged as a hydrogen bonding, and a dot indicates that time. Each unfolding simulation continued till all H-bonds were broken: (a) ST2 simulation = 7 ns, (b) ST4 = 3 ns, (c) ST6 = 11 ns, (d) ST8 = 10 ns, and (e) ST10 = 8 ns.

The $\boldsymbol{v}_i$ and λ_i values derived from the unfolding MD simulation were designated as $\boldsymbol{v}_i$(450K) and λ_i(450K), respectively. The average in Eq. 1 was taken over all of the snapshots in the five unfolding MD trajectories. It is interesting to compare the conformational motions between the 300-K simulation and the unfolding MD simulation. Thus, we also calculated $\boldsymbol{v}_i$ and λ_i from the 300K trajectory, which were designated as $\boldsymbol{v}_i$(300K) and λ_i(300K), respectively. The quantity $|\boldsymbol{v}_i(300\mathrm{K})\cdot\boldsymbol{v}_j(450\mathrm{K})|$ represents the degree of similarity of $\boldsymbol{v}_i$(300K) and $\boldsymbol{v}_j$(450K): if this quantity is 1, the conformational distributions projected onto $\boldsymbol{v}_i$(300K) and $\boldsymbol{v}_j$(450K) are the same, and if it is 0, there is no similarity. In other words, the conformational motions along $\boldsymbol{v}_i$ at 300 K are similar to those along $\boldsymbol{v}_j$ at 450 K, if $|\boldsymbol{v}_i(300\mathrm{K})\cdot\boldsymbol{v}_j(450\mathrm{K})|$ is close to 1. The equation $\boldsymbol{v}_i(450\mathrm{K})\cdot\boldsymbol{v}_j(450\mathrm{K}) = \boldsymbol{v}_i(300\mathrm{K})\cdot\boldsymbol{v}_j(300\mathrm{K})$

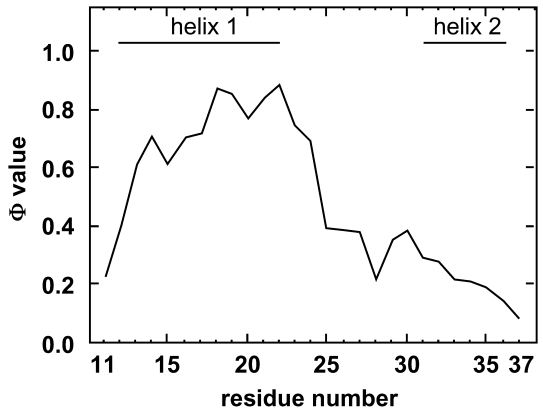


Fig. 2. **Predicted Φ values for C_{trf}.** Helical regions are shown as "helix 1" and "helix 2." The Φ values at residues 12 and 37 are those of the N-terminal cap (*i.e.*, Ace) and the C-terminal cap (*i.e.*, Nme), respectively.

$= \delta_{ij}$ automatically guarantees the following relation: $\Sigma_i [\boldsymbol{v}_i(300K)\cdot\boldsymbol{v}_j(450K)]^2 = \Sigma_j [\boldsymbol{v}_i(300K)\cdot\boldsymbol{v}_j(450K)]^2 = 1$.

RESULTS

Helices 1 and 2 were maintained well around C_{trf} during the 300 K simulation: the main-chain RMSD between the last snapshot in the 300K trajectory and the initial structure was 0.76 Å for helix 1 and that for helix 2 was 1.66 Å. This result ensures that the conformational motion of helix 1 calculated from the 300 K simulation can be used as a control to be compared with that calculated from the unfolding simulations, as mentioned in "MATERIALS AND METHODS." The relative position between helices 1 and 2 changed during the 300 K simulation, which is due to the weakness of the contact between the two helices. Variety of the relative positions between the two helices is also seen in the 20 deposited NMR structures (*12*).

The breaking of H-bonds [HB_1, …, HB_9] in the unfolding simulations is depicted in Fig. 1, which shows the common feature that helix 2 unfolds faster than helix 1, except for HB_8 in ST6. We calculated Φ values for the initial structure, C_{trf}, by a theoretical method developed by Galzitskaya and Finkelstein (*48*). The computed Φ values (Fig. 2) correlate well with the stability of the two helices found in the unfolding MD simulations. The details of the H-bond breaking are discussed below.

The hydrophobic-core formation is frequently observed during the unfolding of C_{trf}. It has been reported that there is no long-range NOE signal in C_{trf} (*12*), and, as mentioned above, the relative positioning between helices 1 and 2 varies in the 20 deposited NMR structures, in each of which a weak contact between helices 1 and 2 is seen. Since the weak contact is formed with the hydrophobic amino-acid residues embedded in the helices, it seems that the contact stabilizes C_{trf}. Figure 3 monitors the radius of gyration (rg_{core}), calculated only from the side-chain atoms of the hydrophobic residues. Here, we denote rg_{core} at C_{trf} as $rg_{core}(C_{trf})$. Figure 4 shows snapshots from the unfolding MD trajectories. A variety of hydrophobic-core formations were found: In the ST2 simulation, rg_{core} fluctuated around $rg_{core}(C_{trf})$ (Fig. 3a), and the conformation was relatively compact with forming

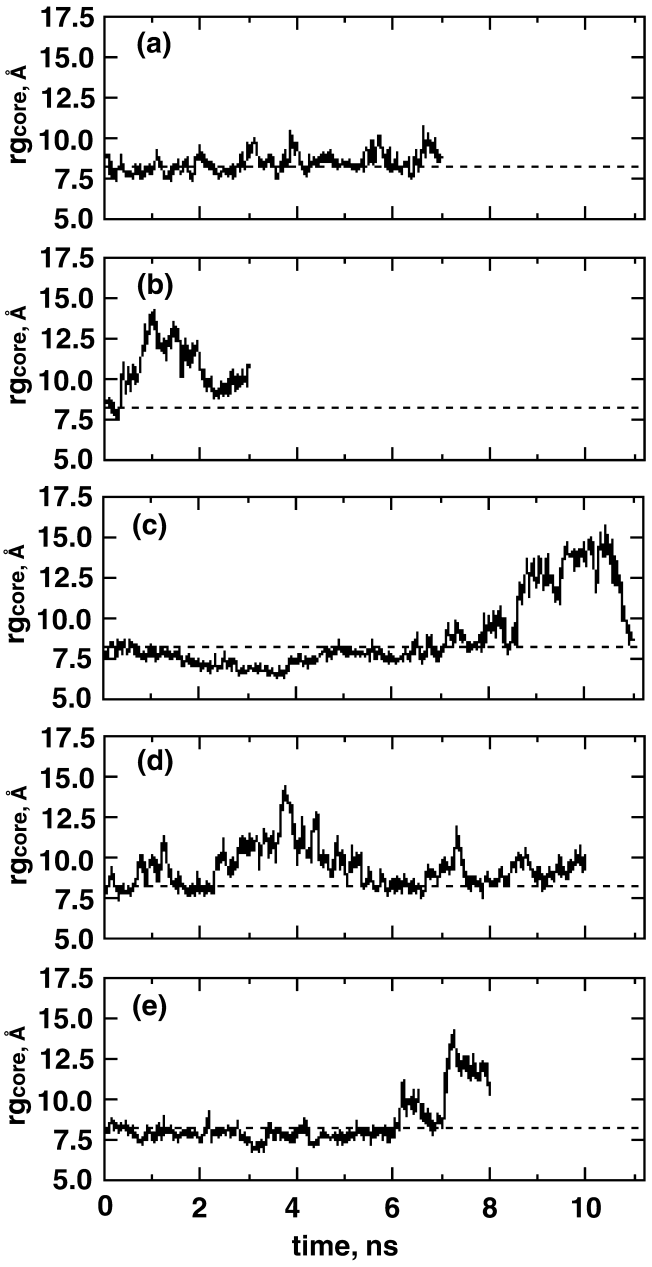
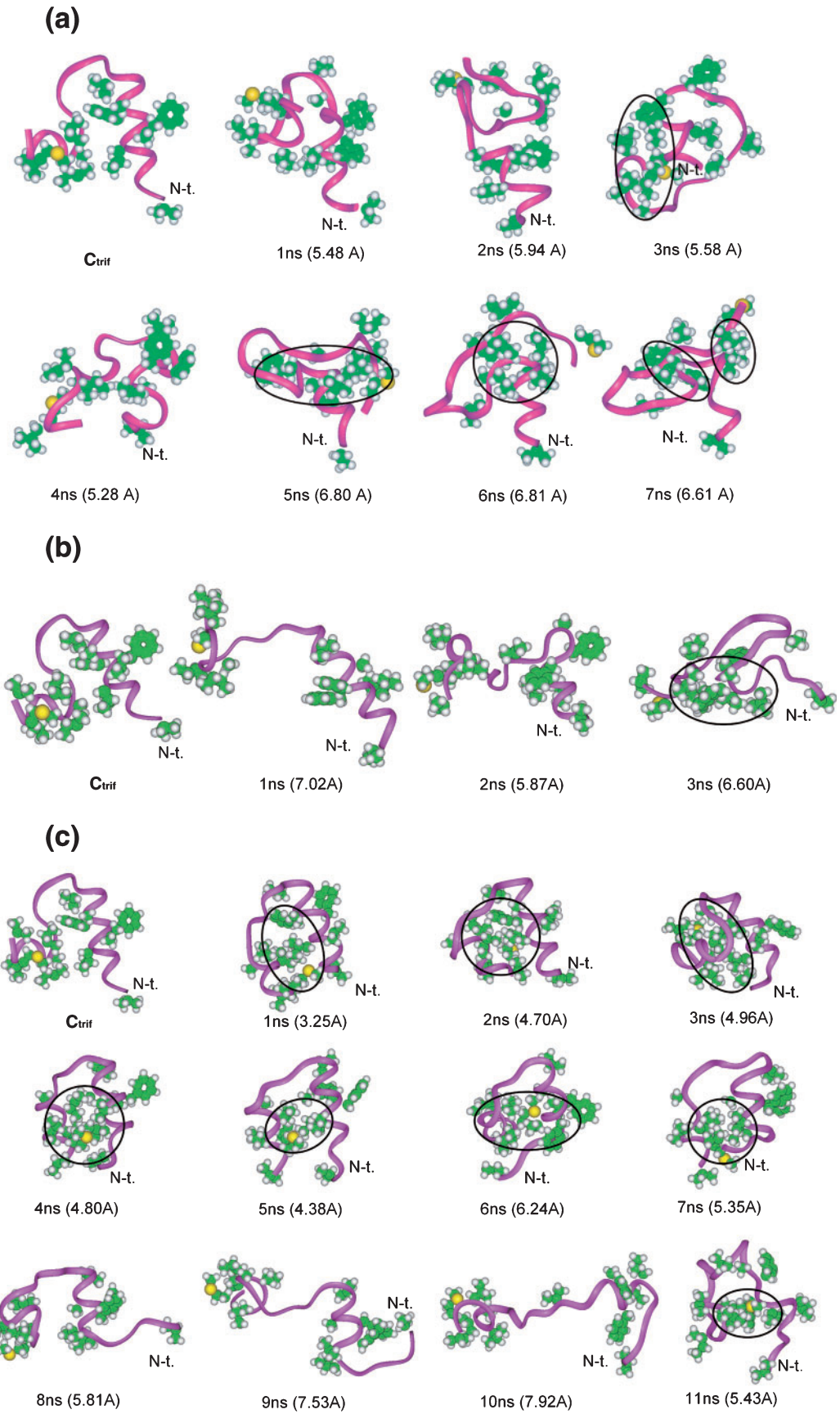


Fig. 3. **Radius of gyration (rg_{core}) of hydrophobic side-chains.** The rg_{core} value was calculated from the side-chain atoms of hydrophobic residues: ^{2}Val, ^{7}Leu-^{11}Ala, ^{14}Val, ^{20}Ala-^{22}Ile, and ^{24}Leu-^{26}Val. Average (shown by broken line) and standard deviations of rg_{core} over snapshots in 300K simulation were 8.256 Å and 0.395 Å, respectively. The average is referred to as $rg_{core}(C_{trf})$ in the text. See legend to Fig. 1 for (a)–(e).

hydrophobic cores (Fig. 4a). In the ST4 simulation, rg_{core} quickly increased and subsequently decreased (Fig. 3b). This behavior can be seen in Fig. 4b: the conformation first expanded, and a hydrophobic core was formed at the end of the simulation. In the ST6 simulation, rg_{core} was smaller than $rg_{core}(C_{trf})$ in the period of 1.5–5 ns (Fig. 3c). Figure 4c shows a hydrophobic-core formation in the period of 1–7 ns, the break-up of the core in the period of 8–10 ns, and another formation at the end of the simulation. The ST8 simulation showed some increase and



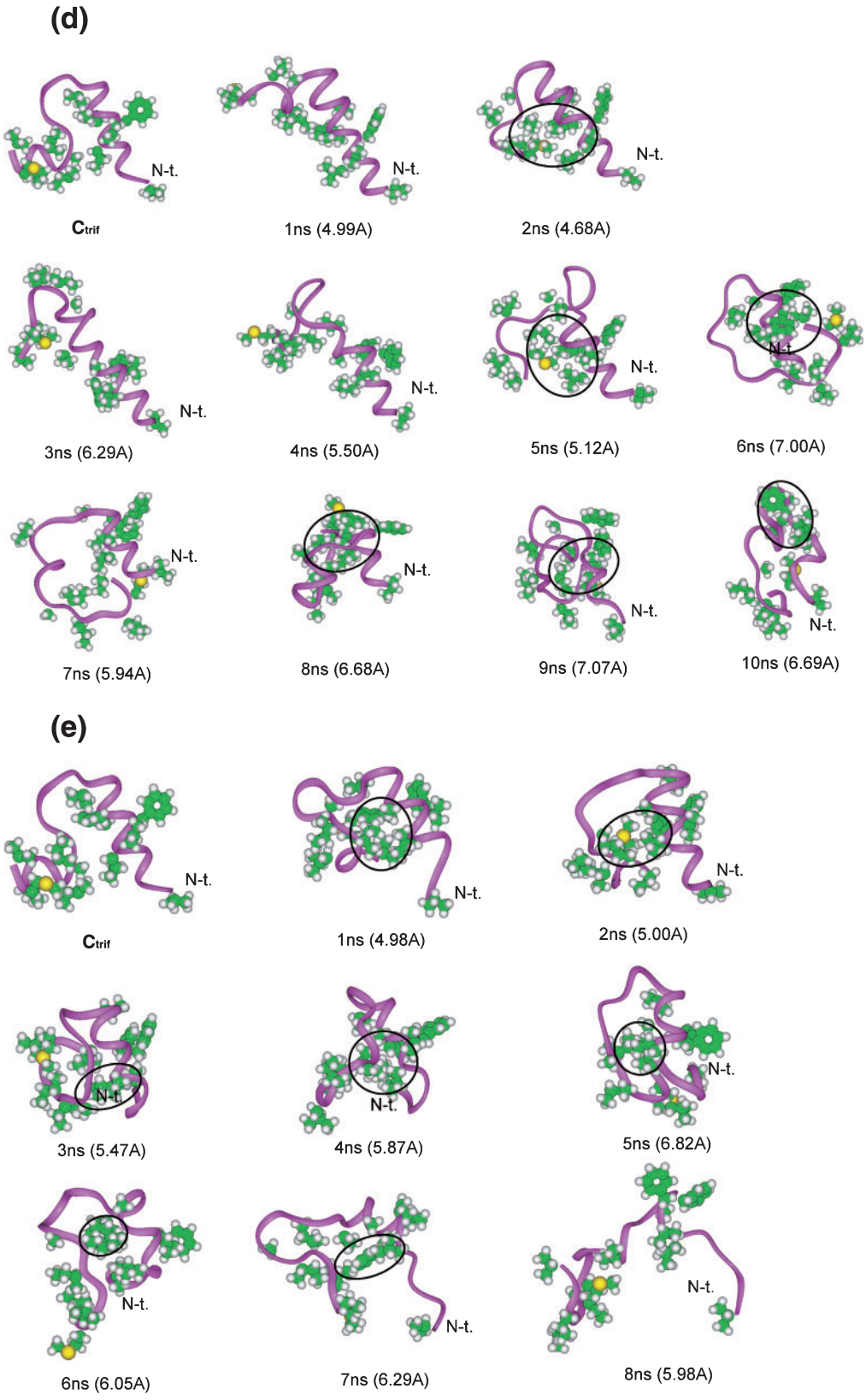


Fig. 4. **Conformations chosen from the unfolding trajectories.** Side-chains of hydrophobic residues are shown by the CPK model with vdw radii smaller than the standard values multiplied by a factor of 0.6. Circles in the solid line represent hydrophobic cores. Values in parentheses near the simulation time are main-chain RMSD between the initial structure (*i.e.*, C_{trf}) and the chosen conformations. "N-t." indicates the position of the N-terminus. See legend to Fig. 1 for (a)–(e).

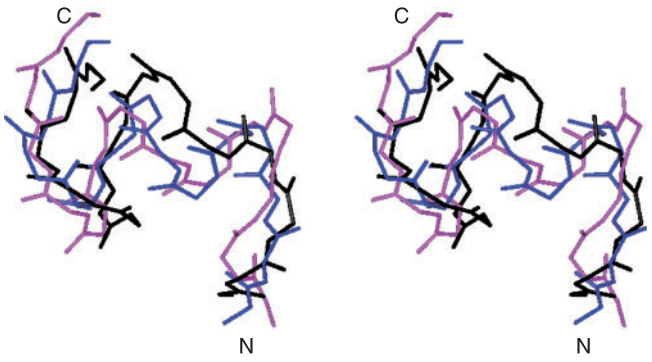


Fig. 5. **Stereo view of sampled conformations similar to C_{wat}.** Shown is a well-converged region, $A\beta_{14-27}$, of the collapsed structure C_{wat}. Model in black is C_{wat}, the one in blue is the conformation at 8,438 ps of the ST6 simulation (main-chain RMSD for $A\beta_{14-27}$ to C_{wat} was 3.355 Å), and the one in magenta is the conformation at 4,344 ps in the ST8 simulation (3.358 Å). "N" and "C" represent the N- and C-termini, respectively.

decrease of rg_{core} (Fig. 3d), and this behavior correlates well with the hydrophobic-core formation (Fig. 4d). Interestingly, prolongation of helix 1 toward the C-terminal side was observed (see conformations at 1, 2, and 3 ns in Fig. 4d). In the ST10 simulation, rg_{core} was slightly smaller than $rg_{core}(C_{trf})$ in the period of 1–6 ns (Fig. 3e). Accordingly, Fig. 4e shows that the conformation in this period was relatively compact with the formation of a hydrophobic core. As shown, the formation of a hydrophobic core was generally found in all of the unfolding MD simulations. However, the amino-acid residues participating in the hydrophobic-core formation varied during the simulation, and, accordingly, the size of the hydrophobic core changed (Fig. 4). Thus, the general feature of the unfolding was the formation and annihilation of various hydrophobic cores.

The collapsed coil structure, C_{wat}, is a stable structure in water as shown by the NMR study (*20*) of $A\beta_{10-35}$ (pdb entry = 1HZ3, where fifteen structures were deposited). Note that the sequence of $A\beta_{10-35}$ is nearly the same as the currently studied one (*i.e.*, $A\beta_{12-36}$). Among the 15 NMR structures, the N- and C-termini (*i.e.*, $A\beta_{10-13}$ and $A\beta_{28-35}$) are structurally diverse, although the main part ($A\beta_{14-27}$) is well converged. Then, we checked if the unfolding MD trajectories involved a similar conformation with $A\beta_{14-27}$ of C_{wat}, and found three similar conformations in the trajectories: the main-chain MRSD of $A\beta_{14-27}$ between them and C_{wat} was smaller than 3.5 Å. Figure 5 shows two conformations of the three. Generally, the unfolding simulation provides conformational changes starting from a low-energy conformation (*i.e.*, a stable conformation or native conformation) to random conformations, because high temperature introduces a strong effect to carry the system toward the denatured state (*i.e.*, the state with a high entropy). Our simulation also provided the unfolding processes: Fig. 1 shows that all of the H-bonds that stabilize helices 1 and 2 were broken in the simulation, and Fig. 3 visually demonstrates a largely fluctuating conformation, which does not converge to C_{wat}. However, Fig. 5 suggests a possibility that the unfolding conformational ensemble may involve a small fraction of conformations relatively close to C_{wat}.

To investigate the details of helix-1 unfolding, PCA was applied to the lengths $[r_1, \ldots, r_7]$. As explained in METHODS, the points $[c_1^1, \ldots, c_7^1]$, $[c_1^2, \ldots, c_7^2]$, …, $[c_1^S, \ldots, c_7^S]$ sampled in the unfolding MD simulation generated a conformational distribution in the 7D space Γ_{7HB}. The standard deviation of the conformational distribution contributed by the i'th axis $\boldsymbol{v}_i$ (*i.e.*, the i'th eigen-vector) was expressed by λ_i, and the whole standard deviation by the summation: $\Sigma_{all}\ \lambda_m(450K) = \lambda_1(450K) + \ldots + \lambda_7(450K)$. The contribution (*i.e.*, $\lambda_i(450K)/\Sigma_{all}\ \lambda_m(450K)$) from each axis to the whole standard deviation was 47.8% for $\boldsymbol{v}_1(450K)$, 21.5% for $\boldsymbol{v}_2(450K)$, 11.4% for $\boldsymbol{v}_3(450K)$, 8.8% for $\boldsymbol{v}_4(450K)$, 5.8% for $\boldsymbol{v}_5(450K)$, 2.8% for $\boldsymbol{v}_6(450K)$, and 1.9% for $\boldsymbol{v}_7(450K)$. Then, the first three eigen-vectors, $\boldsymbol{v}_1(450K)$–$\boldsymbol{v}_3(450K)$, largely contributed to the full standard deviation (*i.e.*, 81%). When the conformation moves only along $\boldsymbol{v}_i$ in Γ_{7HB}, the deviation $\Delta\boldsymbol{R}^i$ of the lengths $[r_1, \ldots, r_7]$ from the mean values $[\langle r_1\rangle, \ldots, \langle r_7\rangle]$ is expressed by the following equation: $\Delta\boldsymbol{R}^i = [r_1 - \langle r_1\rangle, \ldots, r_7 - \langle r_7\rangle] = [\Delta r_1^i, \ldots, \Delta r_7^i]$, where $\Delta r_k^i = \lambda_i(450K)^{1/2} v_{ki}$. A general deviation (*i.e.*, deviation in the unfolding MD simulation) is a mixture (*i.e.*, a linear combination) of individual deviations $\Delta\boldsymbol{R}^i$. Thus, it is important to investigate each deviation. Figure 6 depicts $\Delta\boldsymbol{R}^i$ along each eigen-vector, which indicates the existence of collective modes (*i.e.*, concerted motions) in the unfolding of helix 1: The unfolding along $\boldsymbol{v}_1(450K)$ occurred in such a way that HB_1, HB_2, and HB_3 were stretched with HB_7 being compressed. The unfolding along $\boldsymbol{v}_2(450K)$ pro-

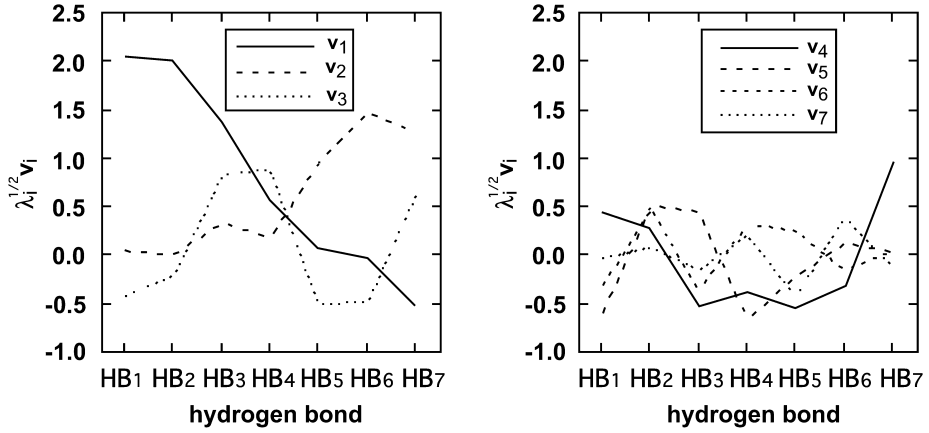


Fig. 6. **Deviation $\Delta\boldsymbol{R}^i$ (*i.e.*, $\boldsymbol{R} - \langle\boldsymbol{R}\rangle$) along $\boldsymbol{v}_i(450K)$ with weight of $\lambda_i(450K)^{1/2}$.**

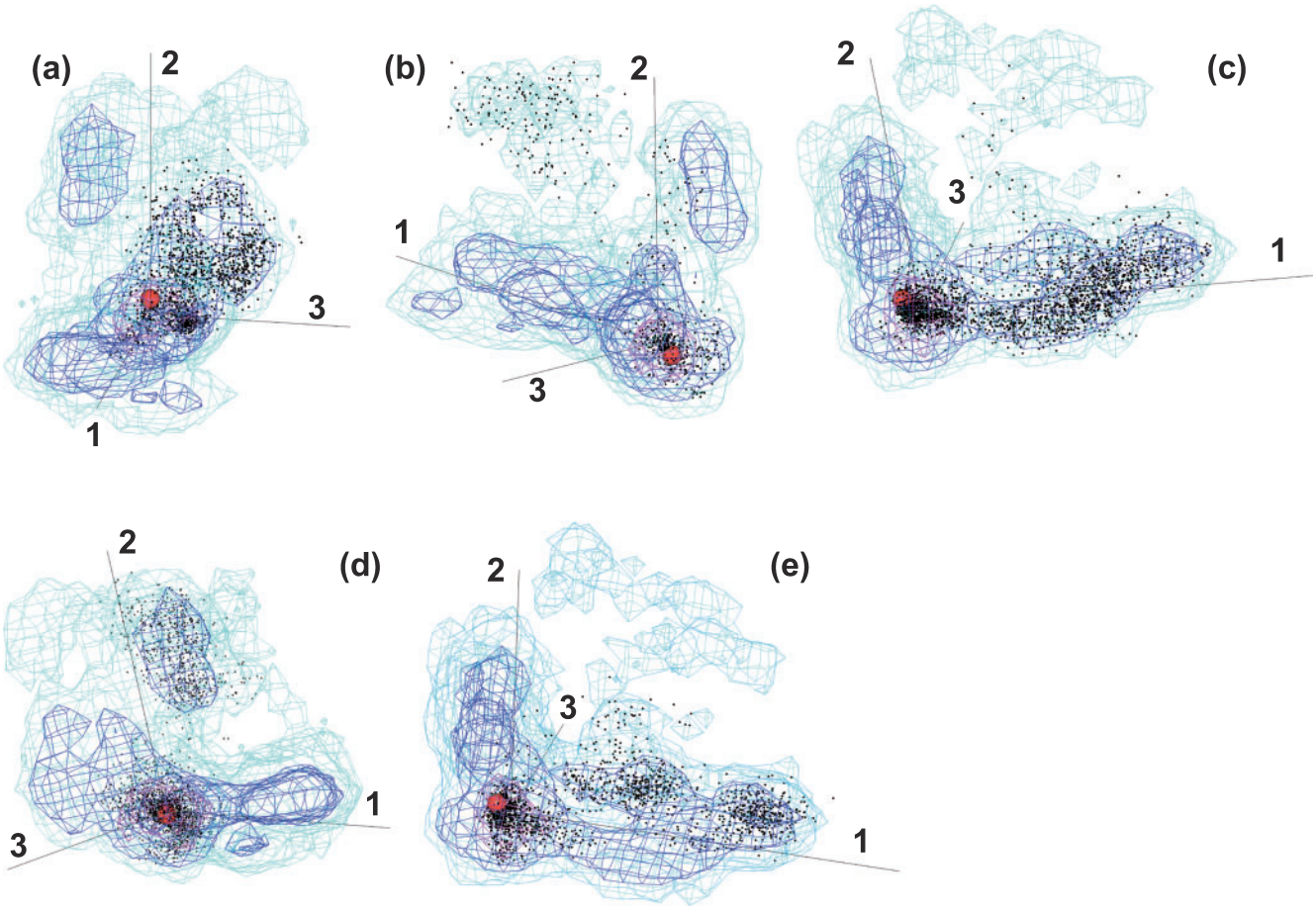


Fig. 7. **Distribution of sampled conformations in sub-space of Γ_{7HB}.** Three axes, $\boldsymbol{v}_1$(450K), $\boldsymbol{v}_2$(450K), and $\boldsymbol{v}_3$(450K), represented by "1," "2," and "3," respectively, construct a 3D sub-space. Distribution is drawn in the form of an iso-density map: the contour in magenta corresponds to density = 0.1 Å^{-3}, the contour in blue to 0.01, and the contour in cyan to 0.001. Black dots are projected conformations picked up from each unfolding MD simulation. The red sphere indicates the position of the initial structure C_{trf}. See legend to Fig. 1 for (a)–(e).

ceeded so that HB_5, HB_6, and HB_7 were stretched with HB_1 being compressed. The unfolding along $\boldsymbol{v}_3$(450K) occurred so that HB_3, HB_4, and HB_7 were stretched with HB_1, HB_5, and HB_6 being compressed. The other eigen-vectors, $\boldsymbol{v}_4$(450K), …, $\boldsymbol{v}_7$(450K), contributed less to the unfolding, because $\Delta\boldsymbol{R}$, contributed by the eigen-vector, became smaller as the eigen-value decreased (see Fig. 6).

The unfolding of helix 1 cannot occur through a single-mode deformation (*i.e.*, a conformational deformation along an eigen-vector), because the breakage of some H-bonds by the eigen-vector accompanies the compression of other H-bonds as shown in Fig. 6. This means that the eigen-vectors should work together to completely unfold helix 1 (*i.e.*, to completely break the H-bonds). The conformations sampled from each unfolding MD simulation were projected on the 3D sub-space contributed by $\boldsymbol{v}_1$(450K), $\boldsymbol{v}_2$(450K), and $\boldsymbol{v}_3$(450K), which yielded the conformational distribution in the sub-space (Fig. 7, a–e). Note that the eigen-vectors were derived from all of the sampled conformations of the five unfolding MD simulations. Here, we describe the scenario of the helix-1 unfolding in each unfolding MD simulation. In the ST2 simulation, the conformations were widely distributed along two axes $\boldsymbol{v}_2$(450K) and $\boldsymbol{v}_3$(450K) (Fig. 7a). As seen in Fig. 1a, HB_6 and HB_7, for which the H-bond stretch was related to $\boldsymbol{v}_2$(450K) shown in Fig. 6, broke at the early stage of the ST2 simulation, and HB_3 and HB_4, for which the H-bond stretch was related to $\boldsymbol{v}_3$(450K), broke in a period of 2–4 ns. On the other hand, the distribution along $\boldsymbol{v}_1$(450K) was narrow. This is because HB_1 and HB_2 were maintained relatively well in the ST2 simulation. In the ST4 simulation, HB_5, HB_6, and HB_7 were broken slightly earlier than the other H-bonds (Fig. 1b). Accordingly, the conformation moved first along $\boldsymbol{v}_2$(450K) after a residing near C_{trf} (Fig. 7b). In the ST6 and ST10 simulations, HB_1, HB_2, and HB_3, for which the H-bond stretch was related to $\boldsymbol{v}_1$(450K), broke first (Fig. 1, c and e). Accordingly, Fig. 7, c and e, demonstrates that the conformation unfolded first along $\boldsymbol{v}_1$(450K). In the ST8 simulation, HB_5, HB_6, and HB_7 broke first (Fig. 1d), and Fig. 7d shows that the unfolding occurred first along $\boldsymbol{v}_2$(450K).

Last, we assessed the similarity of conformational motions between the 300K simulation and the 450 K unfolding simulation. The contribution (*i.e.*, $\lambda_i(300K)/\Sigma_{all}\lambda_m(300K)$) from each axis to the whole standard deviation was 37.3% for $\boldsymbol{v}_1$(300K), 25.6% for $\boldsymbol{v}_2$(300K), 13.8% for

Table 1. $[\boldsymbol{v}_i(300K)\cdot\boldsymbol{v}_j(450K)]^2$.

	$\boldsymbol{v}_1$(300K)	$\boldsymbol{v}_2$(300K)	$\boldsymbol{v}_3$(300K)	$\boldsymbol{v}_4$(300K)	$\boldsymbol{v}_5$(300K)	$\boldsymbol{v}_6$(300K)	$\boldsymbol{v}_7$(300K)
$\boldsymbol{v}_1$(450K)	<u>0.400</u>	0.234	0.036	0.273	0.036	0.015	0.006
$\boldsymbol{v}_2$(450K)	0.235	<u>0.394</u>	0.104	0.096	0.094	0.055	0.021
$\boldsymbol{v}_3$(450K)	0.136	0.004	<u>0.359</u>	0.074	0.069	0.331	0.028
$\boldsymbol{v}_4$(450K)	0.035	0.298	0.200	<u>0.325</u>	0.099	0.003	0.040
$\boldsymbol{v}_5$(450K)	0.126	0.005	0.046	0.066	<u>0.688</u>	0.038	0.031
$\boldsymbol{v}_6$(450K)	0.060	0.059	0.110	0.165	0.006	0.268	<u>0.332</u>
$\boldsymbol{v}_7$(450K)	0.008	0.007	0.145	0.001	0.008	0.290	<u>0.542</u>

The largest element in each row is underlined.

$\boldsymbol{v}_3$(300K), 8.0% for $\boldsymbol{v}_4$(300K), 7.4% for $\boldsymbol{v}_5$(300K), 5.3% for $\boldsymbol{v}_6$(300K), and 2.6% for $\boldsymbol{v}_7$(300K): the contribution of the first three eigen-vectors was about 77%. Thus, the conformational motions from the 300 simulation were mainly determined by the first three collective modes, as well as the case of the unfolding MD simulation. However, the similarity (*i.e.*, 77% and 81%) of the contribution from the first three collective modes between the two simulations does not directly mean that the conformational motions are also similar between the two simulations. The similarity of the conformational motions should be assessed with the quantity $[\boldsymbol{v}_i(300K)\cdot\boldsymbol{v}_j(450K)]^2$, as in Ref. *35*. Table 1 indicates that the diagonal elements $|\boldsymbol{v}_i(300K)\cdot\boldsymbol{v}_i(450K)|^2$ were the largest, except for $i = 6$. This means that the concerted motions in the conformational fluctuations around C_{trf} at 300 K were similar to the unfolding motions at 450 K. However, some off-diagonal elements had non-negligibly large values: $|\boldsymbol{v}_4(300K)\cdot\boldsymbol{v}_1(450K)|^2 = 0.273$, $|\boldsymbol{v}_1(300K)\cdot\boldsymbol{v}_2(450K)|^2 = 0.235$, and $|\boldsymbol{v}_6(300K)\cdot\boldsymbol{v}_3(450K)|^2 = 0.331$. This indicates that some collective modes in $\boldsymbol{v}_i(300K)$ were fused into a concerted motion $\boldsymbol{v}_i(450K)$.

DISCUSSION

Protein folding/unfolding should be discussed based not only on a single simulation, but statistically based on multi-simulations. The unfolding MD simulations demonstrated that helix 2 unfolds first (Fig. 1), as expected from the computed Φ values (Fig. 2). This result is plausible because helix 2 is short and the interaction (*i.e.*, contact) between helix 1 and helix 2 is weak in C_{trf}. The simulation showed that helix 1 unfolds along some pathways as shown in Fig. 7. The N-terminal half of helix 1 had smaller Φ values than the C-terminal half (Fig. 2), which also agrees with the simulation result, because $\boldsymbol{v}_1(450K)$ (*i.e.*, the corrective mode that contributes most to helix-1 unfolding) is related to the stretch (*i.e.*, H-bond break) of HB_1, HB_2, and HB_3 (Fig. 6) located in the N-terminal side of helix 1. Furthermore, the Φ-value computation implies that the C-terminal half of helix 1 may be the folding nucleus in the transition state when Aβ folds from a random conformation to C_{trf}. Thus, this result is a good example to be compared to the free-energy landscape, which will be obtained from a multicanonical MD simulation in future. We emphasize here that the Φ-value computation requires only the initial structure C_{trf} (*48*): *i.e.*, the MD trajectory is not required. Thus, the obtained Φ values have no relation to the MD simulations.

A simplified protein model, the Go-like model (*49*), shows that small globular proteins generally take multiple folding/unfolding pathways in the conformational space (*39*, *50*). The currently observed multiple unfolding pathways of $A\beta_{12-36}$ agree with the results of Go-like model simulations (*39*, *50*). However, the reason for the multiple pathways should be considered individually for each protein. We assume that the multiple pathways of $A\beta_{12-36}$ may be due to the simplicity of the MD initial structure, C_{trf}: Two weakly-interacting helices exposed to solvent are linked by a coil, because the unfolding can start at various positions without a strong influence on the other positions, in spite of the coupled, closely positioned H-bonds (Fig. 6).

A general advantage of MD simulation is the ability to trace the atomic motions at an atomic resolution. In the current study, the simulation highlighted individual events in the unfolding: the patterns of H-bond breaking (Fig. 1), the collective modes in the unfolding (Figs. 6 and 7), the formation and rearrangement of hydrophobic cores (Figs. 3 and 4), the prolongation of helix 1 before complete unfolding (Fig. 4d), and the observation of conformations similar to C_{wat} (Fig. 5).

The current study, based on PCA, provides a method for the quantitative analysis of conformational motions in unfolding processes, and the motions are characterized by collective modes, $\boldsymbol{v}_i(450K)$. Since the simulation was continued until all of the H-bonds were broken, and the life-time of helix 1 was longer than that of helix 2, the helix-1 region partly maintained the helical structure in the MD trajectories. As shown previously, the protein main-chain fold roughly determines the collective motions when the protein conformation fluctuates around a structure (*35*). Therefore, the similarity between $\boldsymbol{v}_i(300K)$ and $\boldsymbol{v}_j(450K)$ shown in Table 1 may be due to the structural similarity between helix 1 in C_{trf} and the partly maintained helix 1 in the unfolding simulation. Protein folding (or unfolding) is a process of conformational changes connecting the native and unfolded states, and in the final stage of folding (or in the initial stage of unfolding), the protein conformation fluctuates near the native structure. This may indicate that some collective modes coupled to each other are used to fold (or unfold) the protein structure.

Multicanonical MD simulation (*51*) enables an exhaustive conformational sampling and provides a free-energy landscape of polypeptides in water at any temperature. Recently we performed the multicanonical MD simulation of $A\beta_{12-36}$ (manuscript in preparation), and have found that $A\beta_{12-36}$ can initiate a prolongation of helix 1, although the probability of occurrence is low. Thus, we assume that the prolongation of helix 1 observed in the ST8 simulation is plausible. The multicanonical MD

simulation also showed a frequent hydrophobic-core formation over a wide temperature range, where the amino-acid residues participating in the core formation were quickly rearranged in the unfolded state. Core formation probably makes the structure more stable at 300 K than at 450 K, although the residue rearrangement at 300 K requires a longer time than at 450 K. The results of the current simulation and the multicanonical MD simulation suggest the importance of the hydrophobic formation. As explained in "INTRODUCTION," a stable structure in water is the collapsed-coil structure, C_{wat} (*20*). Thus, the helix-involving structure, C_{trf}, which is stable in trifluoroethanol/water solution, may spontaneously unfold and refold into C_{wat} at 300 K if the simulation is infinitely prolonged at 300 K. We assume that the rearrangement of the hydrophobic core induces a conformational change, and the conformation becomes stable around C_{wat} by chance.

The observation of conformations similar to C_{wat} in the unfolding (Fig. 5) supports the above assumption. In the unfolded state, the conformation should fluctuate in a wide conformational space, and rarely be close to the stable conformation. Note that the finding of conformations similar to C_{wat} does not mean that the current unfolding simulation can be regarded as a refolding simulation from C_{trf} to C_{wat} with the tracing of a low-free-energy pathway. Remember that both C_{wat} and C_{trf} were unfolded at 450 K in the MD simulations. As quantitatively pointed out by Kamiya et al. in the study of a β-hairpin peptide (*52*), the shape of the free-energy landscape strongly depends on temperature: the probability of the existence of random conformations becomes larger as the temperature increases.

This work was supported by the program MCB RAS and by the Russian Science School program (for O.V.G.).

REFERENCES

1. Tan, S.Y. and Pepys, M.B. (1994) Amyloidosis. *Histopathology* **25**, 403–414
2. Yamamoto, S., Yamaguchi, I., Hasegawa, K., Tsutsumi, S., Goto, Y., Gejyo, F., and Naiki, H. (2004) Glycosaminoglycans enhance the trifluoroethanol-induced extension of β_2-microglobulin-related amyloid fibrils at a neutral pH. *J. Am. Soc. Nephrol.* **15**, 126–133
3. Yamaguchi, I., Hasegawa, K., Takahashi, N., Gejyo, F., and Naiki, H. (2001) Apolipoprotein E inhibits the depolymerization of β_2-microglobulin-related amyloid fibrils at a neutral pH. *Biochemistry* **40**, 8499–8507
4. Takahashi, N., Hasegawa, K., Yamaguchi, I., Okada, H., Ueda, T., Gejyo, F., and Naiki, H. (2002) Establishment of a first-order kinetic model of light chain-associated amyloid fibril extension in vitro. *Biochim. Biophys. Acta*. **1601**, 110–120
5. Yamaguchi, I., Hasegawa, K., Naiki, H., Mitsu, T., Matuo, Y., and Gejyo, F. (2001) Extension of A β_2M amyloid fibrils with recombinant human β_2-microglobulin. *Amyloid* **8**, 30–40
6. Klimov, D.K. and Thirumalai, D. (2003) Dissecting the assembly of Aβ16–22 amyloid peptides into antiparallel β sheets. *Structure* **11**, 295–307
7. Higo, J., Ito, N., Kuroda, M., Ono, S., Nakajima, N., and Nakamura, H. (2001) Energy landscape of a peptide consisting of α-helix, 3_{10}-helix, β-turn, β-hairpin, and other disordered conformations. *Protein Sci.* **10**, 1160–1171
8. Ikeda, K., Galzitskaya, O.V., Nakamura, H., and Higo, J. (2003) β-Hairpins, α-helices, and the intermediates among the secondary structures in the energy landscape of a peptide from a distal β-hairpin of SH3 domain. *J. Comput. Chem.* **24**, 310–318
9. Ikeda, K. and Higo, J. (2003) Free-energy landscape of a chameleon sequence in explicit water and its inherent α/β bifacial property. *Protein. Sci.* **12**, 2542–2548
10. Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J., and Johansson, J. (2001) Prediction of amyloid fibril-forming proteins. *J. Biol. Chem.* **276**, 12945–12950
11. Dobson, C.M. (2001) The structure of protein folding and its links with human disease. *Philos. Trans. Biol. Sci.* **356**, 133–145
12. Sticht, H., Bayer, P., Willbold, D., Dames, S., Hilbich, C., Beyreuther, K., Frank, R.W., and Rosch, P. (1995) Structure of amyloid A4-(1–40)-peptide of Alzheimer's disease. *Eur. J. Biochem.* **233**, 293–298
13. Zagorski, M.G. and Barrow, C.J. (1992) NMR studies of amyloid β-peptides: proton assignments, secondary structure, and mechanism of an α-helix–-β-sheet conversion for a homologous, 28-residue, N-terminal fragment. *Biochemistry* **31**, 5621–5631
14. Talafous, J., Marcinowski, K.J., Klopman, G., and Zagorski, M.G. (1994) Solution structure of residues 1–28 of the amyloid β-peptide. *Biochemistry* **33**, 7788–7796
15. Coles, M., Bicknell, W., Watson, A.A., Fairlie, D.P., and Craik, D.J. (1998) Solution structure of amyloid β-peptide(1–40) in a water-micelle environment. Is the membrane-spanning domain where we think it is? *Biochemistry* **37**, 11064–11077
16. Kohno, T., Kobayashi, K., Maeda, T., Sato, K., and Takashima, A. (1996) Three-dimensional structures of the amyloid β peptide (25–35) in membrane-mimicking environment. *Biochemistry* **35**, 16094–16104
17. Poulsen, S.A., Watson, A.A., Fairlie, D.P., and Craik, D.J. (2000) Solution structures in aqueous SDS micelles of two amyloid β peptides of Aβ(1–28) mutated at the α-secretase cleavage site (K16E, K16F). *J. Struct. Biol.* **130**, 142–152
18. Sonnichsen, F.D., Van Eyk, J.E., Hodges, R.S., and Sykes, B.D. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry* **31**, 8790–8798
19. Sticht, H., Willbold, D., Ejchart, A., Rosin-Arbesfeld, R., Yaniv, A., Gazit, A., and Rosch P. (1994) Trifluoroethanol stabilizes a helix-turn-helix motif in equine infectious-anemia-virus transactivator protein. *Eur. J. Biochem.* **225**, 855–861
20. Zhang, S., Iwata, K., Lachenmann, M.J., Peng, J.W., Li, S., Stimson, E.R., Lu, Y.-a., Felix, A.M., Maggio, J.E., and Lee, J.P. (2000) The Alzheimer's peptide Aβ adopts a collapsed coil structure in water. *J. Struct. Biol.* **130**, 130–141
21. Benzinger, T.L., Gregory, D.M., Burkoth, T.S., Miller-Auer, H., Lynn, D.G., Botto, R.E., and Meredith, S.C. (1998) Propagating structure of Alzheimer's β-amyloid(10–35) is parallel β-sheet with residues in exact register. *Proc. Natl Acad. Sci. USA* **95**, 13407–13412
22. Antzutkin, O.N., Balbach, J.J., Leapman, R.D., Rizzo, N.W., Reed, and J., Tycko, R. (2000) Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of β-sheets in Alzheimer's β-amyloid fibrils. *Proc. Natl Acad. Sci. USA* **97**, 13045–13050
23. Ma, B. and Nussinov, R. (2002) Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers (Aβ16–22, Aβ16–35, and Aβ10–35): sequence effects. *Proc. Natl Acad. Sci. USA* **99**, 14126–14131
24. Li, A. and Daggett, V. (1994) Characterization of the transition state of protein unfolding by use of molecular dynamics: chymotrypsin inhibitor 2. *Proc. Natl Acad. Sci. USA* **91**, 10430–10434
25. Wang, L., Duan, Y., Shortle, R., Imperiali, B., and Kollman, P.A. (1999) Study of the stability and unfolding mechanism of BBA1 by molecular dynamics simulations at different temperatures. *Protein Sci.* **8**, 1292–1304
26. Tsai, J., Levitt, M., and Baker, D. (1999) Hierarchy of structure loss in MD simulations of src SH3 domain unfolding. *J. Mol. Biol.* **291**, 215–225

27. Fersht, A.R. (1995) Characterizing transition states in protein folding: an essential step in the puzzle. *Curr. Opin. Struct. Biol.* **5**, 79–84
28. Mayor, U., Johnson, C.M., Daggett, V., and Fersht, A.R. (2000) Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *Proc. Natl Acad. Sci. USA* **97**, 13518–13522
29. Mayor, U., Guydosh, N.R., Johnson, C.M., Grossman, J.G., Sato, S., Jas, G.S., Freund, S.M.V., Alonso, D.O.V., Daggett, V., and Fersht, A.R. (2003) The complete folding pathway of a protein from nanosecond to microseconds. *Nature* **421**, 863–867
30. Morikami, K., Nakai, T., Kidera, A., Saito, M., and Nakamura, H. (1992) PRESTO: a vectorized molecular mechanics program for biopolymers. *Comput. Chem.* **16**, 243–248
31. Kollman, P.A., Dixon, R.W., Cornell, W.D., Chipot, C., and Pohorille, A. (1997) The development/application of a 'minimalist' organic/biochemical molecular mechanic force field using a combination of *ab initio* calculations and experimental data in *Computer simulations of biological systems* (van Gunsteren, W.F., Weiner, and P.K., Wilkinson, A.J., eds.) Vol. **3**, pp. 83–96, KLUWER/ESCOM, The Netherlands
32. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impley, R.W., and Klein, M.L. (1987) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935
33. Ono, S., Nakajima, N., Higo, J., and Nakamura, H. (2000) Peptide free-energy landscape is strongly dependent on the force field: Comparison of C96 and AMBER95. *J. Comput. Chem.* **21**, 784–762
34. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Kenneth, J., Merz, M., Ferguson D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Amer. Chem. Soc.* **117**, 5179–5197
35. Higo, J. and Umeyama, H. (1997) Protein dynamics determined by backbone conformation and atom packing. *Protein Eng.* **10**, 373–380
36. Ding, H.-Q., Karasawa, N., and Goddard III, W.A. (1992) Atomic level simulations on a million particles: the cell multipole method for Coulomb and London nonbond interactions. *J. Chem. Phys.* **97**, 4309–4315
37. Evans, D.J. and Morriss, G.P. (1983) The isothermal/isobaric molecular dynamics ensemble. *Phys. Lett. A* **98**, 433–436
38. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977) Numerical integration of Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341
39. Koga, N., Takada, S. (2001) Roles of native topology and chain-length scaling in protein folding: A simulation study with a Go-like model. *J. Mol. Biol.* **313**, 171–180
40. Teeter, M.M. and Case, D.A. (1990) Harmonic and quasiharmonic descriptions of crambin. *J. Phys. Chem.* **94**, 8091–8097
41. Hayward, S., Kitao, A., Hirata, F., and Go, N. (1993) Effect of solvent on collective motions in globular protein. *J. Mol. Biol.* **234**, 1207–1217
42. Kitao, A., Hirata, F., and Go, N. (1991) The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum. *Chem. Phys.* **158**, 447–472
43. Amadei, A., Linssen, A.B., and Berendsen, H.J. (1993) Essential dynamics of proteins. *Proteins* **17**, 412–425
44. van Aalten, D.M., Amadei, A., Linssen, A.B., Eijsink, V.G., Vriend, G., and Berendsen, H.J. (1995) The essential dynamics of thermolysin: confirmation of the hinge-bending motion and comparison of simulations in vacuum and water. *Proteins* **22**, 45–54
45. Amadei, A., Linssen. A.B., de Groot, B.L., van Aalten, D.M., and Berendsen, H.J. (1996) An efficient method for sampling the essential subspace of proteins. *J. Biomol. Struct. Dyn.* **13**, 615–625
46. Takahashi, K. and Go, N. (1993) Conformational classification of short backbone fragments in globular proteins and its use for coding backbone conformations. *Biophys. Chem.* **47**, 163–178
47. de Groot, B.L., Daura, X., Mark, A.E., and Grubmuller, H. (2001) Essential dynamics of reversible peptide folding: memory-free conformational dynamics governed by internal hydrogen bonds. *J. Mol. Biol.* **309**, 299–313
48. Galzitskaya, O.V. and Finkelstein, A.V. (1999) A theoretical search for folding/unfolding nuclei in three-dimensional protein structures. *Proc. Natl Acad. Sci. USA* **96**, 11299–11304
49. Go, N. (1983) Theoretical studies of protein folding. *Annu. Rev. Biophys. Bioeng.*, **12**, 183–210
50. Onuchic, J.N., Luthey-Schulten, Z., and Wolynes, , P.G. (1997) Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600
51. Nakajima, N., Nakamura, H., and Kidera, A. (1997) Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides. *J. Phys. Chem.* **B101**, 817–824
52. Kamiya, N., Higo, J., and Nakamura, H. (2002) Conformational transition states of a β-hairpin peptide between the ordered and disordered conformations in explicit water. *Protein Sci.* **11**, 2297–2307