

Helix-Hairpin Transitions of a Designed Peptide Studied by a Generalized-Ensemble Simulation

Satoru G. Itoh,^{*,†} Atsuo Tamura,[‡] and Yuko Okamoto^{*,†}

*Department of Physics, School of Science,
Nagoya University, Nagoya, Aichi 464-8602, Japan, and
Department of Chemistry, Graduate School of Science,
Kobe University, Kobe 657-8501, Japan*

Received November 10, 2009

Abstract: It was recently reported that the designed peptide, whose sequence is INYWLAHAKAGYIVHWTa, has both of the two structures, α -helix and β -hairpin, in aqueous solution. However, the detailed transformation between these two structures is still unclear. In order to study this transformation, we applied a generalized-ensemble simulation to the designed peptide in aqueous solution and deduced the pathways of the designed peptide between α -helix structures and β -hairpin structures.

1. Introduction

Many proteins transform their tertiary structures to carry out their proper biological functions in vivo. Transformations of secondary structures play a role in such tertiary-structure-transformation processes. Therefore, it is important to understand the detailed transformations of the secondary structures. However, it is difficult to investigate the detailed transformations in experiments. Computer simulations are now widely used for such studies.

In order to investigate protein structures and their structure changes, we must realize effective samplings in the conformational space. In the conventional canonical-ensemble simulations,^{1–6} however, it is difficult to achieve this in complex systems such as proteins. This is because the usual canonical-ensemble simulations tend to get trapped in a few of the many local-minimum-energy states. To overcome these difficulties, the generalized-ensemble algorithms have been proposed (for reviews, see, for instance, refs 7 and 8).

We recently developed a new generalized-ensemble algorithm, which is referred to as the multicanonical–multioverlap

algorithm.^{9,10} This algorithm combines the advantages of the multicanonical^{11–14} and multioverlap algorithms.^{15–17} The multicanonical algorithm is one of the most well-known generalized-ensemble algorithms and realizes effective samplings in the conformational space. The multioverlap algorithm also samples effectively the vicinity of specific conformations. The multicanonical–multioverlap algorithm realizes effective samplings in the conformational space more than these two algorithms.^{9,10} It is useful to apply this algorithm to protein systems in order to investigate detailed protein-structure changes.¹⁸

One of the present authors recently designed a new peptide, whose sequence is INYWLAHAKAGYIVHWTa, in order to understand stabilizing mechanisms of protein structures.¹⁹ This peptide has both an α -helix structure (PDB ID code 2DX3) and a β -sheet structure (PDB ID code 2DX4) in aqueous solution. However, the detailed transformation between these two structures is still unclear. Therefore, we applied the multicanonical–multioverlap algorithm to this designed peptide to study the transformation between α -helix and β -sheet structures.

In section 2, we summarize the multicanonical–multioverlap algorithm. In section 3, we describe the details of the multicanonical–multioverlap molecular dynamics (MD) simulations that we performed. We present their results in section 4. Section 5 is devoted to conclusions.

2. Methods

2.1. Multicanonical–Multioverlap Algorithm. In the multicanonical–multioverlap algorithm,^{9,10} by employing the non-Boltzmann weight factor W_{mco} , which we refer to as the multicanonical–multioverlap weight factor, a uniform probability distribution with respect to the potential energy and a dihedral-angle distance is obtained:

$$P_{\text{mco}}(E, d) = n(E, d)W_{\text{mco}}(E, d) \equiv \text{constant} \quad (1)$$

where E is the potential energy and d is a dihedral-angle distance. The dihedral-angle distance is defined by

$$d = \frac{1}{n\pi} \sum_{i=1}^n d(v_i, v_i^0) \quad (2)$$

where n is the total number of dihedral angles, v_i is the dihedral angle i , and v_i^0 is the dihedral angle i of the reference conformation. The distance $d(v_i, v_i^0)$ between two dihedral angles is given by

$$d(v_i, v_i^0) = \min(|v_i - v_i^0|, 2\pi - |v_i - v_i^0|) \quad (3)$$

* Corresponding author e-mail: itoh@tb.phys.nagoya-u.ac.jp (S.G.I.), okamoto@phys.nagoya-u.ac.jp (Y.O.).

[†] Nagoya University.

[‡] Kobe University.

The multicanonical-multioverlap weight factor at a constant temperature T_0 can be written as

$$W_{\text{mco}}(E, d) = e^{-\beta_0 E_{\text{mco}}(E, d)} \quad (4)$$

where β_0 is defined by $\beta_0 = 1/k_B T_0$ (k_B is the Boltzmann constant) and $E_{\text{mco}}(E, d)$ is the multicanonical-multioverlap potential energy. Equation 1 implies that multicanonical-multioverlap simulations realize a random walk in the two-dimensional potential-energy and dihedral-angle-distance space and are able to effectively sample the conformational space. Accordingly, we can obtain accurate free-energy landscapes of protein systems and estimate folding pathways and the transition states among the specific conformations.^{9,10} We remark that we employed the MD version of the multicanonical-multioverlap algorithm¹⁸ in this article.

2.2. Reweighting Techniques. The results from a multicanonical-multioverlap simulation can be analyzed by the reweighting techniques. Suppose that we have determined the multicanonical-multioverlap potential energy E_{mco} in eq 4 at a constant temperature T_0 and that we performed the simulation at this temperature. The expectation value of a physical quantity A at any temperature T is calculated from

$$\langle A \rangle_T = \frac{\sum_{E, d} A(E, d) n(E, d) e^{-\beta E}}{\sum_{E, d} n(E, d) e^{-\beta E}} \quad (5)$$

where the best estimate of the density of states is given by the single-histogram reweighting techniques^{20,21} (see eq 1):

$$n(E, d) = \frac{N_{\text{mco}}(E, d)}{W_{\text{mco}}(E, d)} \quad (6)$$

and $N_{\text{mco}}(E, d)$ is the histogram of the probability distribution that was obtained by the multicanonical-multioverlap simulation.

We can also calculate the free energy (or the potential of mean force) with appropriate reaction coordinates. For example, the free energy $F(\xi_1, \xi_2; T)$ with reaction coordinates ξ_1 and ξ_2 at temperature T is given by

$$F(\xi_1, \xi_2; T) = -k_B T \ln P_c(\xi_1, \xi_2; T) \quad (7)$$

where $P_c(\xi_1, \xi_2; T)$ is the reweighted canonical probability distribution of ξ_1 and ξ_2 and given by (see eq 5)

$$P_c(\xi_1, \xi_2; T) = \frac{\sum_{E, d} N_{\text{mco}}(\xi_1, \xi_2; E, d) e^{\beta_0 E_{\text{mco}}(E, d) - \beta E}}{\sum_{\xi_1, \xi_2, E, d} N_{\text{mco}}(\xi_1, \xi_2; E, d) e^{\beta_0 E_{\text{mco}}(E, d) - \beta E}} \quad (8)$$

and $N_{\text{mco}}(\xi_1, \xi_2; E, d)$ is the histogram of the probability distribution that was obtained from the multicanonical-multioverlap simulation.

3. Computational Details

It was reported that the designed peptide, whose sequence is INYWLAHAKAGYIVHWTA, has the two characteristic structures, α -helix and β -hairpin, coexisting in a pH 4.5 solution in experiments.¹⁹ In our simulation, the histidine residues of the



Figure 1. Reference conformation in our multicanonical-multioverlap simulation. The N terminus and the C terminus are on the top side and on the bottom side, respectively. The figure was created with RasMol.²⁸

Table 1. Backbone Dihedral Angles of the Reference Conformation in Figure 1

angle (deg)		angle (deg)	
ϕ_2	-153.7	ϕ_{10}	-76.4
ψ_2	153.0	ψ_{10}	-30.0
ϕ_3	-96.8	ϕ_{11}	-95.4
ψ_3	-51.5	ψ_{11}	15.5
ϕ_4	-67.6	ϕ_{12}	-70.7
ψ_4	-44.6	ψ_{12}	-3.9
ϕ_5	-83.9	ϕ_{13}	-69.8
ψ_5	22.6	ψ_{13}	-8.9
ϕ_6	-125.9	ϕ_{14}	-64.5
ψ_6	-56.3	ψ_{14}	-46.4
ϕ_7	-83.3	ϕ_{15}	-84.4
ψ_7	-29.0	ψ_{15}	-45.6
ϕ_8	-85.3	ϕ_{16}	-64.1
ψ_8	-1.4	ψ_{16}	-83.8
ϕ_9	-101.3	ϕ_{17}	-161.2
ψ_9	-50.2	ψ_{17}	143.1

designed peptide were protonated in order to conform our simulation conditions to the low pH ones in the experiment. The force field that we adopted is the CHARMM 22 parameter set.²² We employed the Generalized Born/Surface Area (GB/SA) model^{23–25} as an implicit solvent model.

In multicanonical-multioverlap simulations, we must have a reference conformation. We adopted the conformation in Figure 1 as the reference conformation in this article. This conformation was obtained by minimizing the α -helix structure of the designed peptide which corresponds to the model 1 structure in the 2DX3 PDB file.¹⁹ We list the backbone-dihedral angles of the reference conformation in Table 1, and these values were employed as the reference dihedral angles ν_i^0 of the dihedral-angle distance in eq 2. Hence, we took into account only the backbone-dihedral angles ϕ (the rotation angles around the N–C $_{\alpha}$ bonds) and ψ (the rotation angles around the C $_{\alpha}$ –C bonds) of the residues 2–17 of the designed peptide as the reference dihedral angles in our simulations. We remark that, although we employed the α -helix structure as a reference

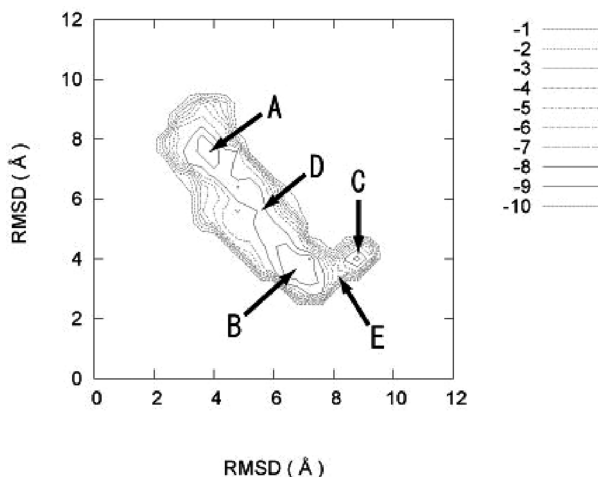


Figure 2. Free-energy landscape obtained from the results of the multicanonical–multioverlap MD simulation at 300 K. The abscissa and the ordinate are the RMSDs of the backbone C_α atoms with respect to the α -helix structure in Figure 1 and the hairpin structure in Figure 3, respectively. Contour lines are drawn every 1 kcal/mol. The labels A, B, and C locate the free-energy local-minimum states. The labels D and E stand for the saddle points between A and B and between B and C, respectively.

conformation, results of multicanonical–multioverlap simulations are independent of selection of reference conformations as long as sampling is performed sufficiently.^{9,10}

The multicanonical–multioverlap weight factor was first determined so that a free random walk was realized in the two-dimensional energy-overlap space. As for the potential-energy random walk, this weight factor covered the temperature range from 300 to 600 K (see Supporting Information for the determination of the multicanonical–multioverlap weight factor). We then performed a multicanonical–multioverlap production run, in which the time step was 0.5 fs, for 44.5 ns after an equilibration of 0.5 ns. The initial conformation of the designed peptide for the production run was a random-coil conformation.

4. Results and Discussion

We show the free-energy landscape at 300 K in Figure 2. The free-energy landscape was obtained from the results of the multicanonical–multioverlap MD simulation by the reweighting techniques in eqs 7 and 8. The abscissa is the root-mean-square distance (RMSD) of the backbone C_α atoms with respect to the reference conformation in Figure 1. The RMSD with respect to the reference conformation is defined by

$$\text{RMSD} = \min \left[\sqrt{\frac{1}{N} \sum_i (\mathbf{q}_i - \mathbf{q}_i^0)^2} \right] \quad (9)$$

where N is the number of atoms, $\{\mathbf{q}_i^0\}$ are the coordinates of the reference conformation, and the minimization is over the rigid translations and rigid rotations of the coordinates of the conformation $\{\mathbf{q}_i\}$ with respect to the center of geometry. The ordinate is the RMSD with respect to the hairpin structure of the designed peptide in Figure 3. This structure was obtained by minimizing the hairpin structure of the model 1 structure in the 2DX4 PDB file.¹⁹ Three free-energy local-minimum states and two transition states among these local-minimum states are

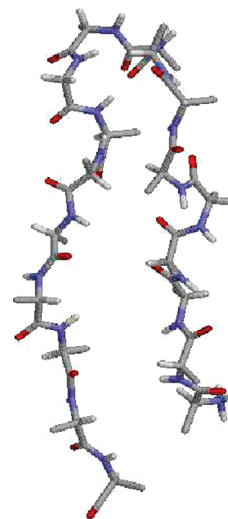


Figure 3. The hairpin structure of the designed peptide. The N-terminus and the C-terminus are on the right-hand side and on the left-hand side, respectively. The figure was created with RasMol.²⁸

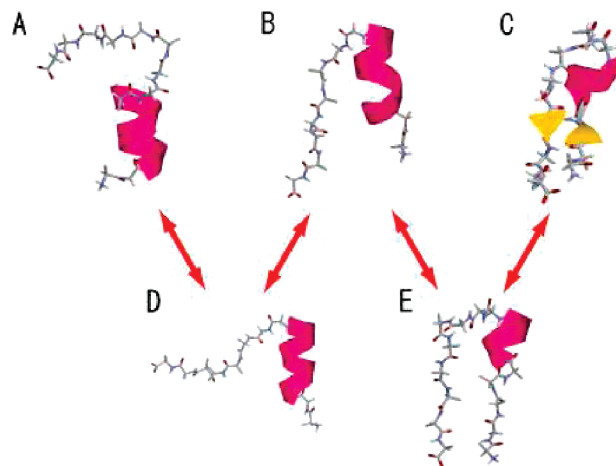


Figure 4. Typical structures in the corresponding local-minimum states in Figure 2. The N-terminus and the C-terminus are on the right-hand side and on the left-hand side, respectively. The arrows indicate possible pathways of the designed peptide from α -helix structure to hairpin-like structure. The figures were created with RasMol.²⁸

identified in Figure 2. We label the three local-minimum states as A, B, and C; the transition state between the local-minimum states A and B as D; and the transition state between the local-minimum states B and C as E. These local-minimum states were not completely coincident with the experimental conformations in Figures 1 and 3. In other words, RMSDs between the local-minimum states and the experimental conformations were more than 2 Å at 300 K. This is because conformations whose RMSDs were less than 2 Å had high potential energy with the CHARMM force field. We remark that the free-energy landscape with respect to the dihedral-angle distance d is described in the Supporting Information as a reference.

In Figure 4, we show typical conformations of the designed peptide in each local-minimum state (see the Supporting Information for how these typical conformations were obtained). In the free-energy local-minimum state A, the designed peptide

Table 2. Expectation Values of the α -Helix Content and the End-to-End Distance and Values of Free Energy in the Local-Minimum States in Figure 2^a

	A	B	C	Figure 1	Figure 3
α -helicity (%)	27.1 (2.2)	29.5 (5.0)	22.0 (0.3)	38.9	0.0
end-to-end distance (Å)	23.8 (0.8)	13.2 (0.9)	4.5 (0.3)	25.7	7.0
free energy (kcal/mol)	−9.2 (0.4)	−8.9 (0.3)	−8.5 (1.5)		

^a The corresponding values for the conformations in Figures 1 and 3 are also given as reference. Errors were estimated by the jackknife method.^{29–31}

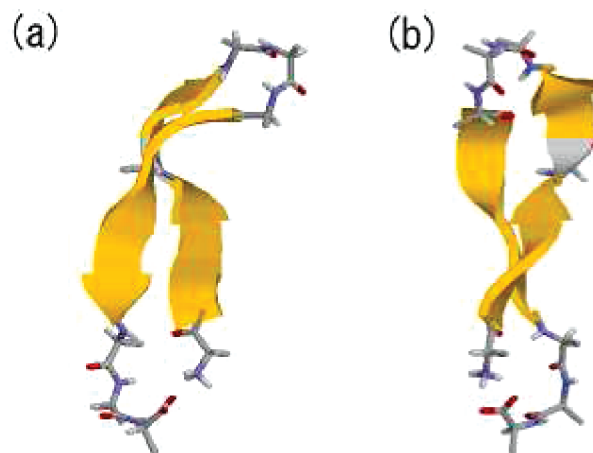
has α -helix structures with an extended C-terminus. In B, this extended C-terminus is close to the N-terminus. The extended C-terminus forms a β -ladder with the N-terminus in the local-minimum state C. These hairpin-like structures are similar to the structure in Figure 3. We also show typical conformations in transition states in Figure 4. In the transition state D between the free-energy local-minimum states A and B, the extended C-terminus is almost vertical with respect to the helix axis. Because the C-terminal parts are extended to the upper side of the α -helix structure in A and to the lower side of the α -helix structure in B such as the conformations in Figure 4, the conformation in the transition state D is very reasonable as an intermediate conformation between those in A and B. The structures in the transition state E between the local-minimum states B and C have an extended N-terminus, and this extended N-terminus is created by breaking the corresponding part of the α -helix structure in B. We remark that, although we determined the conformations in the transition states in Figure 4 simply as shown in the Supporting Information, there is a useful analysis method to obtain more accurate transition states among free-energy local minima that was presented in ref 26.

We list in Table 2 the expectation values of the α -helicity and end-to-end distance and the values of the free energy in the local-minimum states in Figure 2. Moreover, we listed the α -helicity and end-to-end distance of the conformations in Figures 1 and 3 as references. Here, an α -helicity of a conformation is defined by

$$\frac{N_{\alpha}}{N_{\text{res}}} \quad (10)$$

where N_{α} is the number of residues that have helix structures and N_{res} is the number of the residues of the peptide ($N_{\text{res}} = 18$ for the designed peptide). We employed the definition of secondary structure of proteins (DSSP) criteria²⁷ to determine the secondary structures of the peptide. The α -helicity in the local-minimum state C is small in comparison with those in the local-minimum states A and B from this table. This is because the structures in C have a short extended N-terminus by breaking apart of N-terminal α -helix structures in A and B as mentioned above. Furthermore, the average end-to-end distance gets smaller when the state is transferred from A to B or from B to C. The expectation value of the end-to-end distance in the local-minimum state C is close to the end-to-end distance obtained from the experimental conformation in Figure 3.

From Figures 2 and 4 and Table 2, we deduce the transformation process of secondary structures of the designed peptide between α -helix structures and hairpin-like structures

**Figure 5.** A β -hairpin structure in the free-energy local-minimum state C in Figure 2. a and b correspond to the same conformation viewed from different angles. The figures were created with RasMol.²⁸

as follows. Stage 1: The designed peptides have an α -helix structure with an extended C-terminus such as the structure in A in Figure 4. Stage 2: The extended C-terminus is bent (transition state D) and comes close to the N-terminus like the structure in B. Stage 3: A part of the α -helix on the N-terminus is broken, and the extended N-terminus is formed (transition state E). Stage 4: The extended C-terminus and the extended N-terminus create the β -ladder such as the structure in the local-minimum state C. These pathways are summarized in Figure 4 (see the arrows).

The β -hairpin structures of the designed peptide were observed in an experiment.¹⁹ Although the typical conformations in the local-minimum did not have complete β -hairpin structures in our simulation, complete β -hairpin structures did exist during the simulation, where the conformations were in C. Figure 5 shows the complete β -hairpin structure found in the free-energy local-minimum state C. This structure is formed by breaking the remaining α -helix structure of the typical conformation in C.

5. Conclusions

It is important to understand the transformations of the secondary structures of proteins. This is because the protein functions are closely associated with the tertiary structures, and the tertiary structure changes are caused by the secondary structure changes. The typical secondary structures for the proteins are α -helix and β -sheet structures, and the transformations between α -helix and β -sheet structures are often important for the protein function. For instance, amyloidogenesis is caused by the transformations from α -helix to β -sheet structures for certain proteins.

In the present work, we calculated the free-energy landscape at 300 K for the designed peptide in aqueous solution from the results of a multicanonical–multioverlap MD simulation. The designed peptide has the two structures, α -helix and β -hairpin, coexisting in an experiment. We observed these structures in our simulation and identified intermediate structures and transition states between the two structures. We also deduced the transformation pathways of the designed peptide between α -helix structures and β -hairpin structures. In other words, it

was clarified that the α -helix structure of the N-terminus is broken step by step, and the β -ladder is created between the broken parts and the extended C-terminus in the transformations from α -helix to β -hairpin structures. Therefore, we believe that our results for the transformations between α -helix and β -sheet structures will play an important role in understanding the protein functions, while our results are still with a small designed peptide.

Acknowledgment. The computations were performed on the computers at the Research Center for Computational Science, Institute for Molecular Science. This work was supported, in part, by the Grants-in-Aid for the Next Generation Super Computing Project, Nanoscience Program and for Scientific Research on Innovative Areas, "Fluctuations and Biological Functions," from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Supporting Information Available: Computational details. Free-energy landscape with respect to the dihedral-angle distance. This information is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087.
- (2) Hoover, W. G.; Ladd, A. J. C.; Moran, B. High-strain-rate plastic-flow studied via non-equilibrium molecular-dynamics. *Phys. Rev. Lett.* **1982**, *48*, 1818.
- (3) Evans, D. J. Computer experiment for non-linear thermodynamics of couette-flow. *J. Chem. Phys.* **1983**, *78*, 3297.
- (4) Nosé, S. A molecular-dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255.
- (5) Nosé, S. A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511.
- (6) Hoover, W. G. Canonical dynamics - equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695.
- (7) Mitsutake, A.; Sugita, Y.; Okamoto, Y. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers* **2001**, *60*, 96.
- (8) Itoh, S. G.; Okumura, H.; Okamoto, Y. Generalized-ensemble algorithms for molecular dynamics simulations. *Mol. Sim.* **2007**, *33*, 47.
- (9) Itoh, S. G.; Okamoto, Y. A new generalized-ensemble algorithm: multicanonical-multioverlap algorithm. *Mol. Sim.* **2007**, *33*, 83.
- (10) Itoh, S. G.; Okamoto, Y. Effective sampling in the configurational space of a small peptide by the multicanonical-multioverlap algorithm. *Phys. Rev. E* **2007**, *76*, 026705.
- (11) Berg, B. A.; Neuhaus, T. Multicanonical algorithms for 1st order phase-transitions. *Phys. Lett. B* **1991**, *267*, 249.
- (12) Berg, B. A.; Neuhaus, T. Multicanonical ensemble - a new approach to simulate 1st-order phase-transitions. *Phys. Rev. Lett.* **1992**, *68*, 9.
- (13) Hansmann, U. H. E.; Okamoto, Y.; Eisenmenger, F. Molecular dynamics, Langevin and hybrid Monte Carlo simulations in a multicanonical ensemble. *Chem. Phys. Lett.* **1996**, *259*, 321.
- (14) Nakajima, N.; Nakamura, H.; Kidera, A. Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides. *J. Phys. Chem. B* **1997**, *101*, 817.
- (15) Berg, B. A.; Noguchi, H.; Okamoto, Y. Multioverlap simulations for transitions between reference configurations. *Phys. Rev. E* **2003**, *68*, 036126.
- (16) Itoh, S. G.; Okamoto, Y. Multi-overlap molecular dynamics methods for biomolecular systems. *Chem. Phys. Lett.* **2004**, *400*, 308.
- (17) Itoh, S. G.; Okamoto, Y. Theoretical studies of transition states by the multioverlap molecular dynamics methods. *J. Chem. Phys.* **2006**, *124*, 104103.
- (18) Itoh, S. G.; Okamoto, Y. Amyloid- β (29–42) dimer formations studied by a multicanonical-multioverlap molecular dynamics simulation. *J. Phys. Chem. B* **2008**, *112*, 2767.
- (19) Araki, M.; Tamura, A. Transformation of an α -helix peptide into a β -hairpin induced by addition of a fragment results in creation of a coexisting state. *Proteins* **2007**, *66*, 860.
- (20) Ferrenberg, A. M.; Swendsen, R. H. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* **1988**, *61*, 2635.
- (21) Ferrenberg, A. M.; Swendsen, R. H. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* **1989**, *63*, 1658.
- (22) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586.
- (23) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
- (24) Dominy, B. N.; Brooks, C. L., III. Development of a generalized born model parametrization for proteins and nucleic acids. *J. Phys. Chem. B* **1999**, *103*, 3765.
- (25) Feig, M.; Brooks, C. L., III. Evaluating CASP4 predictions with physical energy functions. *Proteins* **2002**, *49*, 232.
- (26) Bolhuis, P. G.; Dellago, C.; Chandler, D. Reaction coordinates of biomolecular isomerization. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5877.
- (27) Kabsch, W.; Sander, C. Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577.
- (28) Sayle, R. A.; Milner-White, E. J. Biomolecular Graphics for All. *Trends Biochem. Sci.* **1995**, *20*, 374.
- (29) Quenouille, M. H. Notes on bias in estimation. *Biometrika* **1956**, *43*, 353.
- (30) Miller, R. G. Jackknife - review. *Biometrika* **1974**, *61*, 1.
- (31) Berg, B. A. *Markov Chain Monte Carlo Simulations and Their Statistical Analysis*; World Scientific: Singapore, 2004.

CT9005932