Absence of reptation in the high-temperature folding of the trpzip2 β-hairpin peptide

Jed W. Pitera[a)]
IBM Almaden Research Center, San Jose, California 95120

Imran Haque
University of California at Berkeley, Berkeley, California 94720

William C. Swope
IBM Almaden Research Center, San Jose, California 95120

(Received 7 February 2006; accepted 3 March 2006; published online 14 April 2006)

We have carried out extensive all atom explicit solvent simulations of the high-temperature folding and unfolding of the trpzip2 β-hairpin peptide and examined the resulting trajectories for evidence of folding via a reptation mechanism. Over 300 microcanonical simulations of 10 ns each were initiated from a Boltzmann ensemble of conformations at 425 K. Though we observed numerous folding and unfolding events, no evidence of reptation was found. The diffusional dynamics of the peptide are orders of magnitude faster than any observed reptation-like motion. Our data suggest that the dominant mechanisms for β-hairpin folding under these conditions are hydrophobic collapse and turn formation, and that rearrangements occur via significant expansion of the polypeptide chain. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2190226]

INTRODUCTION

While the general physical principles of protein folding are well understood,[1] there is still a great deal of controversy about the specific mechanism of folding. A particularly well-studied example is the process of β-hairpin folding. There are a number of peptides that fold to stable β-hairpins in solution.[2] Researchers have identified at least three putative mechanisms for folding in these systems: turn formation followed by the sequential formation of hydrogen bonds from the turn towards the ends of each strand ("zipping"),[3] hydrophobic collapse followed by rearrangement;[4] and a reptation-like mechanism, where a misregistered state with non-native backbone hydrogen bonds rearranges to the native structure by a sliding motion of the backbone along its length.[5] The original "zipping" mechanism was proposed on the basis of temperature-jump fluorescence experiments,[3] but many simulations have shown either a hydrophobic collapse mechanism or a hybrid of the two, with concurrent collapse and hydrogen bond formation.[6] The observed mechanism also depends somewhat on the definitions used for hydrogen bonds and hydrophobic contacts.[7] The reptation mechanism was first identified in simulation studies[5] using a simplified model of the protein coupled with a systematic transition state search method, the activation-relaxation technique.[8] In addition to reptation, this systematic search found examples of zipping and collapse mechanisms. A recent Monte Carlo study[9] also reported evidence of reptation in a simplified β-hairpin model. Though significant populations of misregistered states have been observed in β-hairpin simulations with all-atom potentials,[10] the reptation mechanism has not been observed in these cases.

Reptation is a dynamical motion that polymers undergo in confined or entangled environments.[11] It consists of a sliding motion of the polymer along its length in a narrow tube of free volume. This sliding mechanism implies a mean-squared displacement of the monomers that is proportional to $t^{1/4}$ and an overall center of mass diffusion constant proportional to N^{-2}. Typical β-hairpin peptide systems range from nine[12] to sixteen[13] amino acids, and are studied in dilute solution. Given these short chain lengths, reptation behavior is unexpected, and we set out to search for examples of reptation-like β-hairpin folding using a detailed all-atom model in explicit solvent.

METHODS AND RESULTS

The simulated system consisted of the unblocked trpzip2 peptide (SWTWENGKWTWK), 1 Cl^- counterion, and 3605 water molecules (11034 atoms). The solute was modeled with the AMBER parm96 force field[14] and the solvent with the TIP3P[15] water model. All simulations were carried out using the PMEMD module of AMBER8.[16] Each kinetic trajectory was an independent microcanonical (NVE) simulation, using a 2 femtosecond (fs) timestep with SHAKE[17] to constrain all bonds with a tolerance of 10^{-6}. The particle mesh Ewald method was used for long-range electrostatics with an alpha parameter of 0.30768, a grid spacing of 1 Å, and 4th order interpolation. A cutoff of 9.0 Å was applied to the van der Waals and direct electrostatic interactions.

Initial coordinates for each of 327 trajectories were taken from a replica-exchange molecular dynamics (REMD) simulation[18] of trpzip2 that had been started in a fully ex-

[a)]Author to whom correspondence should be addressed. Telephone: 408-927-2084. Fax: 408-927-2100. Electronic mail: pitera@us.ibm.com

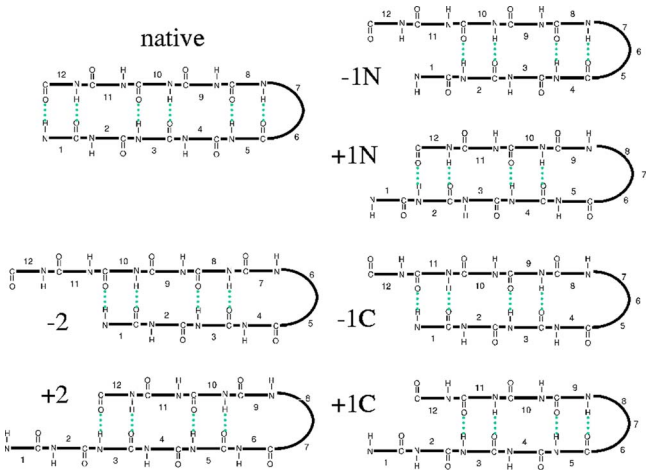

FIG. 1. Schematics of the native and misregistered states of trpzip2 examined in this work. Misregistered states are identified by the direction of displacement of the two strands (±), the degree of displacement (1 or 2 residues) and N or C to indicate a rotation about the N or C terminal strand of the hairpin at the turn.

tended conformation. The initial structure for the REMD simulation had been equilibrated at 310 K and 1 atm to yield a final cubic simulation cell 48.095 Å on a side. Eighty replicas were used to span a temperature range from 245 to 600 K. An initial 26.8 nanosecond (ns) REMD run was sufficient to yield statistically equivalent distributions at 425 K from REMD calculations started in both the folded and unfolded states. Although it is above the experimental melting temperature of 345 K,[19] the peptide is still 5–10% folded at 425 K in the REMD simulations. Starting coordinates for the kinetic trajectories were sampled every 40 picoseconds (ps) from a subsequent 13 ns of REMD simulation. Each simulation was initialized with velocities from a Maxwell-Boltzmann distribution at 425 K, center of mass motion removed, and then simulated for 10 ns. Coordinates and energies were saved every 1 ps for subsequent analysis. The final data set of 327 10 ns trajectories corresponds to an aggregate simulation time of 3.27 μs, and includes over 3 million conformations.

We calculated the C_α-RMSD of each conformation from the trpzip2 NMR structure (PDB[20] ID 1LE1).[19] Based on a histogram of these values,[23] we counted as folded any structure within 2.0 Å C_α-RMSD of the NMR structure, while any structure more than 3.0 Å C_α-RMSD was considered unfolded. There were 45 trajectories with at least one folding event (transition from the unfolded to the folded state) and 67 trajectories with at least one unfolding event. In 35 of the folding trajectories, the peptide was extremely unfolded (>9 Å C_α-RMSD) at some point before folding. Of the initial structures, 35 were folded and 280 unfolded by this metric, with two falling between the two C_α-RMSD thresholds.

Structural analyses were also carried out to assign each conformation a native, misregistered, or unfolded state using hydrogen bond criteria. Schematics of the native and misregistered states are shown in Fig. 1. Backbone hydrogen bonds were identified via geometric criteria (a donor-acceptor distance of <4.0 Å and a donor-H-acceptor angle >120 deg) and secondary structures were assigned with the STRIDE algorithm.[21] Native and misregistered states were identified based on the patterns listed in Table I. Conformations were counted as unfolded if they had two or fewer backbone hydrogen bonds. Several misregistered states were significantly populated, but none of our trajectories contained direct transitions between native and misregistered states as would be expected for reptation (Table II). The presence of significantly stable structured but non-native states in the unfolded ensemble has recently been seen in several other beta-hairpin simulations.[10] Native and misregistered states both rearrange by passing through the unfolded state. In addition, the different misregistered states only interconvert by passing through the unfolded state (data not shown). An analysis based on secondary structure rather than hydrogen bonds produced similar results.

TABLE I. Hydrogen bond patterns, secondary structure, and populations observed for the native structure and misregistered states of the trpzip2 β-hairpin. In the secondary structure pattern, "E" represents β-strand, "T" a turn and "x" any secondary structure. Conformations were assigned to a state based on the hydrogen bond pattern.

State	Hydrogen bonded residue pattern	Corresponding secondary structure motif (Ref. 21)	Number of conformations observed
Native	1-12,3-10,5-8	EEEEETTEEEEE	220 093
−2	1-10,3-8	EEExxxxEEExx	31
+2	3-12,5-10	xxEEExxxxEEE	22
−1N	2-10,4-8	xEEExxxEEExx	1 286
+1N	2-12,4-10	xEEExxxxxEEE	4 151
−1C	1-11,3-9	EEExxxxxEEEx	0
+1C	3-11,5-9	xxEEExxxEEEx	17 415

The previous reports of reptation[5] described a mechanism which involves the near-simultaneous breakage of four or more non-native hydrogen bonds and formation of four or more native hydrogen bonds. The shortest time required to break four non-native hydrogen bonds and form four native hydrogen bonds in our simulations was 180 ps, and this process on average takes 5.6 ns. The reverse process, breaking four native hydrogen bonds and forming four non-native hydrogen bonds, takes an average of 5.5 ns and a minimum of 110 ps.

We also looked for reptation with other geometric criteria. We searched for evidence of a "sliding" motion of the central turn but did not find any. When the chain reptates in a forward direction, monomer i moves closer to the position previously occupied by monomer $i+1$, monomer $i+1$ to the position occupied by monomer $i+2$, and so on. Conse-

TABLE II. Self- and non-self transitions observed between native, misregistered, and unfolded states in the trpzip2 simulations. Unfolded states are defined as having two or fewer intramolecular backbone hydrogen bonds.

	Destination		
Origin	Native	Misregistered	Unfolded
Native	206 896	0	431
Misregistered	0	19 200	53
Unfolded	426	64	242 022

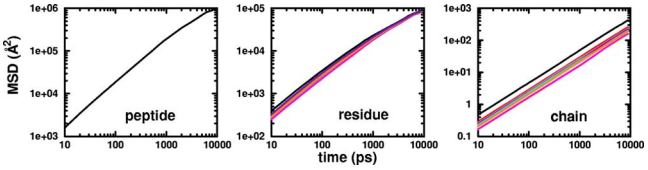

FIG. 2. Mean-squared displacements (in Å^2) of the center of mass of the peptide ("peptide"), the center of mass of each residue ("residue"), and the center of mass of each residue projected onto vectors connecting successive centers of mass ("chain").

quently, we calculated the "shifted C_α-dRMS" for shift of ±1 or 2 for different conformations in each trajectory and compared it to the C_α-dRMS between those conformations. No reptation was observed, and the only pairs of conformations where the shifted C_α-dRMS was smaller were cases where both structures were helical and alternate ends of the helix had unfolded. The details of these and other geometric analyses are provided in the supplementary material.[23]

We also calculated the mean-squared displacement for the center of mass of the peptide and center of mass of each residue as a function of lag time. Figure 2 shows a plot of both these quantities. The peptide center of mass diffusion constant is 0.003 25 cm^2/sec, while the individual residue center of mass displacements have a sublinear time dependence, ranging from $t^{0.87}$ to $t^{0.93}$. This is clearly a stronger time dependence than the $t^{1/4}$ expected for a reptation mechanism. To monitor reptation, we also calculated the mean-squared displacement of each residue projected onto a "chain coordinate" defined by the vectors connecting successive residue centers of mass. If reptation is a dominant mechanism for rearrangement of the peptide, movement along the chain should account for a significant fraction of each residue's mean-squared displacement, but this is clearly not the case.

DISCUSSION AND CONCLUSIONS

We did not observe folding or unfolding via a reptation mechanism in a large data set of over 300 high-temperature microcanonical simulations of the β-hairpin peptide trpzip2 that totals 3.27 μs of simulation time. Both structural and dynamical observables were monitored for signs of reptation.

There are several factors that disfavor reptation-like mechanisms in protein folding. First, reptation is fundamentally driven by confinement. It describes the motions of polymer chains in dense, entangled melts, where other modes of conformational change are excluded by steric constraints. In contrast, a protein in solution is relatively unconstrained, and has many more possible avenues for reorganization. The narrow "tube" of backbone positions allowed by a reptation mechanism defines a very low entropy state, compared to the many alternative conformations the protein can adopt in solution. The high temperature used in these simulations should therefore disfavor the reptation mechanism, and it remains to be seen if similar calculations at a temperature closer to the melting temperature might show reptation. A second factor arises from the chemical nature of the polypeptide backbone. Residues that participate in a tight type I′ turn (ASP6 and GLY7 in trpzip2) have ϕ torsions greater than zero. For residues other than glycine, transitions to this region of the Ramachandran plot are rare, and involve surmounting a large energy barrier.[22] As a result, the protein chain cannot simply "slide past" the turn region. It is possible that a longer β-hairpin sequence with a more open loop might exhibit reptation, though the observed misregistered states of trpzip2 have more open loops than the native fold and still were not seen to reptate.

While it is impossible to prove the absence of a reptation mechanism, we feel that our simulations provide a compelling argument that reptation is not a primary mechanism of conformational rearrangement in trpzip2 at high temperature, and perhaps other β-hairpin peptides under other conditions. We did not observe reptation by any of several criteria, even over a large number of folding, misfolding, and unfolding trajectories. Instead, the dynamics of trpzip2 at high temperature are Rouse-type, as expected for a polymer in dilute solution, and conformational transitions largely proceed via highly unfolded states.

ACKNOWLEDGMENTS

All of the simulations described in this work were carried out on the MareNostrum supercomputer at the Barcelona Supercomputing Center, Universidad Politécnica de Catalunya, Barcelona, Spain. WCS and IH acknowledge support from NSF MRSEC Center on Polymer Interfaces and Macromolecular Assemblies DMR-0213618.

[1] A. R. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (W.H. Freeman, 1998).
[2] S. H. Gellman, Curr. Opin. Chem. Biol. **2**, 717 (1998).
[3] V. Munoz, P. A. Thompson, J. Hofrichter, and W. A. Eaton, Nature (London) **390**, 196 (1997).
[4] A. R. Dinner, T. Lazaridis, and M. Karplus, Proc. Natl. Acad. Sci. U.S.A. **96**, 9068 (1999).
[5] G. H. Wei, N. Mousseau, and P. Derreumaux, Proteins-Structure Function and Bioinformatics 56, 464 (2004); G. H. Wei, P. Derreumaux, and N. Mousseau, J. Chem. Phys. **119**, 6403 (2003).
[6] R. H. Zhou, B. J. Berne, and R. Germain, Proc. Natl. Acad. Sci. U.S.A. **98**, 14931 (2001).
[7] W. C. Swope, J. W. Pitera, F. Suits, M. Pitman, M. Eleftheriou, B. G. Fitch, R. S. Germain, A. Rayshubski, T. J. C. Ward, Y. Zhestkov, and R. Zhou, J. Phys. Chem. **108**, 6582 (2004).
[8] G. T. Barkema and N. Mousseau, Phys. Rev. Lett. **77**, 4358 (1996).
[9] H. Imamura and J. Z. Y. Chen, Proteins-Structure Function and Bioinformatics (in press) (published online, 2006).
[10] S. V. Krivov and M. Karplus, Proc. Natl. Acad. Sci. U.S.A. **101**, 14766 (2004); X. W. Wu and B. R. Brooks, Biophys. J. **86**, 1946 (2004).
[11] P. G. de Gennes, J. Chem. Phys. **55**, 572 (1971).
[12] F. J. Blanco, M. A. Jimenez, J. Herranz, M. Rico, J. Santoro, and J. L. Nieto, J. Am. Chem. Soc. **115**, 5887 (1993).
[13] M. S. Searle, D. H. Williams, and L. C. Packman, Nat. Struct. Biol. **2**, 999 (1995).
[14] W. D. Cornell, J. W. Caldwell, and P. A. Kollman, Journal De Chimie Physique Et De Physico-Chimie Biologique 94, 1417 (1997).
[15] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[16] D. A. Case, AMBER6 (University of California, San Francisco, 2004).
[17] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[18] U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997); Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[19] A. G. Cochran, N. J. Skelton, and M. A. Starovasnik, Proc. Natl. Acad. Sci. U.S.A. **98**, 5578 (2001).
[20] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, Nucleic Acids Research 28, 235 (2000).

[21] D. Frishman and P. Argos, Proteins-Structure Function and Genetics 23, 566 (1995).
[22] M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, and R. A. Friesner, J. Am. Chem. Soc. **119**, 5908 (1997).
[23] See EPAPS Document No. E-JCPSA6-124-541613 for the histogram of observed C_α-RMSD values as well as further details of the mean-squared displacement calculations and geometric analyses of reptation. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).