Folding and Translocation of the Undecamer of Poly-L-leucine across the Water−Hexane Interface. A Molecular Dynamics Study

Christophe Chipot[†,‡] and Andrew Pohorille*[,‡,§]

Contribution from the Laboratoire de Chimie Théorique, Unité de Recherche Associée au CNRS No. 510, Université Henri Poincaré-Nancy I, BP. 239, 54506 Vandœuvre-lès-Nancy Cedex, France, Exobiology Branch, NASA−Ames Research Center, MS 239-4, Moffett Field, California 94035-1000, and Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143-0446

Received January 2, 1998. Revised Manuscript Received April 28, 1998

Abstract: The undecamer of poly-L-leucine at the water−hexane interface is studied by molecular dynamics simulations. This represents a simple model relevant to folding and insertion of hydrophobic peptides into membranes. The peptide, initially placed in a random coil conformation on the aqueous side of the system, rapidly translocates toward the hexane phase and undergoes interfacial folding into an α-helix in the subsequent 36 ns. Folding is nonsequential and highly dynamic. The initially formed helical segment at the N-terminus of the undecamer becomes transiently broken and, subsequently, reforms before the remainder of the peptide folds from the C-terminus. The formation of intramolecular hydrogen bonds during the folding of the peptide is preceded by a dehydration of the participating polar groups, as they become immersed in hexane. Folding proceeds through a short-lived intermediate, a 3_{10}-helix, which rapidly interconverts to an α-helix. Both helices contribute to the equilibrium ensemble of folded structures. The helical peptide is largely buried in hexane, yet remains adsorbed at the interface. Its preferred orientation is parallel to the interface, although the perpendicular arrangement with the N-terminus immersed in hexane is only slightly less favorable. In contrast, the reversed orientation is highly unfavorable, because it would require dehydration of C-terminus carbonyl groups that do not participate in intramolecular hydrogen bonding. For the same reason, the transfer of the undecamer from the interface to the bulk hexane is also unfavorable. The results suggest that hydrophobic peptides fold in the interfacial region and, simultaneously, translocate into the nonpolar side of the interface. It is further implied that peptide insertion into the membrane is accomplished by rotating from the parallel to the perpendicular orientation, most likely in such a way that the N-terminus penetrates the bilayer.

1. Introduction. Background and Significance

In order to perform cellular functions, many proteins have to translocate across the water−membrane interface. These include a wide variety of receptors, channels, and transporters, which are integral to the membrane,[1] and water-soluble proteins, which cross membranes to reach their final destination.[2,3] Most proteins adsorb to the membrane surface in nonnative states. The subsequent insertion into the lipid bilayer is often spontaneous[4] and requires extensive conformational rearrangements. In integral membrane proteins, these rearrangements typically involve the formation of folded, most often α-helical, segments which typically assemble into large, multisubunit aggregates. Proteins that translocate through membranes anchor to the bilayer surfaces *via* signal sequences, many of which fold into α-helices in the membrane environment.[2,3] Helical fragments in many integral proteins and bacterial signal sequences share a common feature: they exhibit a highly hydrophobic character.

Since the incorporation of proteins into membranes is a complex process, it is common to turn to model peptides that insert into membranes and remain helical in the bilayer environment.[5−8] The conceptual basis for this simplification is formed by the assumption that the formation of helical segments is an early event, essentially independent of other stages of folding or translocation.[9] Insertion into membranes and folding of peptides are also of interest in their own rights, especially for understanding interactions between membranes and naturally occurring toxins, hormones, and membrane fusion peptides.[10] Despite considerable efforts, however, many structural, energetic, and dynamic aspects of these processes are still poorly understood. This is largely due to difficulties in applying the most powerful experimental techniques for studying structure and dynamics of biopolymers, such as X-ray crystallography, multidimensional nuclear magnetic resonance, and stopped-flow spectroscopy, to proteins and peptides in the membrane environment.

In the present work, phenomena related to folding and translocation of hydrophobic peptides across the water−membrane interface are studied *via* molecular dynamics (MD) computer simulations. Among the issues of special interest are

[†] Université Henri Poincaré-Nancy I.
[‡] NASA−Ames Research Center.
[§] University of California, San Francisco.

(1) Gennis, R. B. *Biomembranes: Molecular Structure and Function*; Springer: New York, 1989.
(2) Jones, J. D.; McKnight, J. C.; Gierasch, L. M. *J. Bioenerg. Biomembr.* **1990**, *22*, 213.
(3) Tamm, L. K. *Biochim. Biophys. Acta* **1991**, *1071*, 123.
(4) Maduke, M.; Roise, D. *Science* **1993**, *260*, 364.
(5) Moll, T. S.; Thompson, T. E. *Biochemistry* **1994**, *33*, 15469.
(6) Zhang, Y. P.; Lewis, R. N. A. H.; Hodges, R. S.; McElhaney, R. N. *Biochemistry* **1995**, *34*, 2348.
(7) Zhang, Y. P.; Lewis, R. N. A. H.; Hodges, R. S.; McElhaney, R. N. *Biochemistry* **1995**, *34*, 2362.
(8) Saleh, M. T.; Ferguson, J.; Boggs, J. M.; Gariepy, J. *Biochemistry* **1996**, *35*, 9325.
(9) Popot, J. L.; Engelman, D. M. *Biochemistry* **1990**, *29*, 4031.
(10) Kaiser, E. T.; Kezdy, F. J. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 561.

10.1021/ja980010o CCC: $15.00
Published on Web 11/10/1998

the role of the interface in mediating the folding process, the existence of folding intermediates, the relationship between folding and insertion, and the preferred location and orientation of the peptide relative to the interface.

The model of a hydrophobic peptide selected for the simulation is the Ac− and −NHMe terminally blocked undecamer of poly-L-leucine. The choice of L-leucine was motivated by its propensity to form α-helices,[11] but also by the hydrophobic nature of its side chain, which is anticipated to interact favorably with nonpolar environments, thereby making poly-L-leucine a good candidate for a prototypical transmembrane peptide. Even though this peptide is fairly short, capable of forming only three full turns in an α-helical conformation, simulating its folding at the water−membrane interface is a challenging task. The main anticipated difficulty is related to folding times, which are slow in comparison with time scales characteristic of MD simulations. These are, in part, due to a slow relaxation of the collective motions of lipid molecules in the bilayer. In addition, the incorporation of the peptide into the membrane is associated with the spatial rearrangement of the lipid chains, which must expand in the directions parallel to the bilayer plane. This type of rearrangement cannot be treated correctly in conventional MD. Instead, a modified, extended system treatment is required. Although the conceptual basis for such a treatment has been worked out,[12−14] several technical issues still remain to be resolved, and applications of these methods have been, so far, very limited.

To avoid these difficulties, which might have rendered the simulations of peptide folding at the water−membrane interface impossible, we considered a somewhat simpler system—the water−hexane interface. Clearly, several effects potentially important for interfacial folding and peptide insertion into membranes, such as interactions between lipid head groups and polar groups in the peptide backbone, perturbations of lipid chains in the bilayer and transmembrane potentials,[15,16] are neglected in this simplified system. The water−hexane interface, nevertheless, retains the main characteristic of the water−membrane system: the close proximity between the aqueous phase and the nonpolar environment. In several experimental[5,17−19] and theoretical[20−23] studies, it has been argued that hydrophobic interactions resulting from the coexistence of two phases of different polarity provide the major driving force for peptide insertion into the membrane. Thus, carefully interpreted results on interactions between peptides and the water−hexane interface can be very helpful for understanding the behavior of these peptides at the water−membrane interface.

Different aspects of secondary structure transitions in hydrophobic peptides and the insertion of such peptides into membranes were probed at different levels of approximation in several previous theoretical studies. Unfolding of L-alanine-rich peptides in water was studied in MD simulations with explicit representation of all molecules in the system.[24−27] These simulations, however, were carried out at elevated temperatures[25] or employed restrictions on the unfolding pathways.[26] Folding of similar peptides in an aqueous solution was also considered, but only utilizing a continuum model of the solvent.[28] Both folding and insertion into membranes were investigated by applying the lattice model of peptides and an effective medium approximation for both the water and the membrane phases.[22] Insertion of peptides, already folded into α-helices, was treated by several authors,[23,29,30] all of which used a continuum dielectric representation of the surroundings. Probably the most detailed application of this general approach was the one presented recently by Ben-Tal *et al.*[23] Stability of a hydrophobic helix after its insertion into the membrane was studied in a detailed MD simulation.[31] The present work is distinguished from the previous studies by explicitly addressing the problem of folding and translocation across the interface and treating it in full atomic detail on a realistic time scale, without imposing any artificial constraints on the folding pathway.

2. Method

All the molecular dynamics simulations were carried out in the microcanonical (N, V, E) ensemble, using the program COSMOS.[32] The interfacial environment of the Ac− and −NHMe terminally blocked undecamer of poly-L-leucine consisted of a lamella of 1380 water molecules in contact with a lamella of 409 hexane molecules. Both the water and the hexane lamellae were in equilibrium with their respective vapor phases. The dimensions of the simulation box were 42 × 42 Å in the *x*- and *y*-directions, parallel to the interface, and 200 Å in the *z*-direction, perpendicular to the interface. Periodic boundary conditions were applied in all three directions. The equations of motion were solved using the Verlet algorithm,[33] with a time step of 2 fs. The length of the MD trajectory was 50 ns. To maintain the average temperature at 300 K, the velocities of the particles were occasionally rescaled. This rescale was necessary to correct for inaccurate integration of the equations of motion, and occurred on average every 5000 steps, *i.e.*, every 10 ps. On average, the drift in the total energy between two rescales of the velocities never exceeded 10 kcal/mol, which represents a deviation of *ca.* 0.09% during this time frame. Over 100 ps of MD trajectory, the average deviation in the total energy amounted to *ca.* 0.7%.

The water molecules were described by the TIP4P model,[34] and the hexane molecules were represented by the OPLS potential energy functions.[35] In the model of hexane, $-CH_3$ and $-CH_2-$ groups were described by single, united atoms of appropriate radius. This model of alkanes has proven to be successful for predicting relevant thermodynamic quantities, such as the water−alkane surface tension[36]

(11) O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646.
(12) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.
(13) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117.
(14) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.
(15) Jacobs, R. E.; White, S. H. *Biochemistry* **1989**, *28*, 3421.
(16) Dalbey, R. E.; Kuhn, A.; Von Heijne, G. *Trends Cell Biol.* **1995**, *5*, 380.
(17) Shin, Y. K.; Levinthal, C.; Levinthal, F.; Hubbell, W. L. *Science* **1993**, *259*, 960.
(18) Wang, Y.; Weiner, H. *Biochemistry* **1994**, *33*,12860.
(19) Dathe, M.; Schumann, M.; Wieprecht, T.; Winkler, A.; Beyermann, M.; Krause, E.; Matsuzaki, K.; Murase, O.; Bienert, M. *Biochemistry* **1996**, *35*, 12612.
(20) Jähnig, F. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3691.
(21) Milik, M.; Skolnick, J. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9391.
(22) Milik, M.; Skolnick, J. *Proteins* **1993**, *15*, 15.
(23) Ben-Tal, N.; Ben-Shaul, A.; Nicholls, A.; Honig, B. *Biophys. J.* **1996**, *70*, 1803.
(24) Tobias, D. J.; Broks, C. L., III. *Biochemistry* **1991**, *30*, 6059.
(25) Daggett, V.; Kollman, P. A.; Kuntz, I. D. *Biopolymers* **1991**, *31*, 1115.
(26) Tirado-Rives, J.; Maxwell, D. S.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1993**, *115*, 11590.
(27) Young, W. S.; Brooks, C. L., III. *J. Mol. Biol.* **1996**, *259*, 560.
(28) Snug, S. S.; Wu, X. W. *Proteins* **1996**, *25*, 202.
(29) Edholm, O.; Johansson, J. *Eur. Biophys. J.* **1987**, *14*, 203.
(30) Biggin, P. C.; Sansom, M. S. *Biophys. Chem.* **1996**, *60*, 99.
(31) Shen, L.; Bassolino-Klimas, D.; Stouch, T. R. *Biophys. J.* **1997**, *73*, 3.
(32) Owenson, B.; Pohorille, A. *COSMOS − A software package for COmputer Simulations of MOlecular Systems*, NASA−Ames Research Center, Moffett Field, CA 94035-1000, 1987.
(33) Verlet, L. *Phys. Rev.* **1967**, *159*, 98.
(34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(35) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6683.

and the free energy of solvation of organic molecules.[37,38] Although the united atom model appears to be adequate for treating equilibrium properties of uncharged systems, it is insufficient if charged species are present in the nonpolar phase[39] or if accurate estimates of dynamic properties are sought. In particular, it yields overestimating diffusion rates in the nonpolar phase. It is, thus, probable that the folding times estimated from the united atom model represent lower bounds to the actual times. The nonbonded and the intramolecular parameters in the potential energy function for the peptide were taken from the AMBER all-atom force field of Cornell *et al.*[40] Parameters in the water−solute and the hexane−solute potential energy functions were obtained from the standard OPLS combination rules.[35]

Pairwise intermolecular interactions involving water molecules and/or small, electrically neutral groups of poly-L-leucine and hexane molecules were smoothly truncated between 7.5 and 8.0 Å by means of a cubic-spline switching function[41,42] applied to both the energy and the forces. Thus, long-range, electrostatic interactions between the solute and the solvent were not taken into account. This approximation is justified by the hydrophobic nature of the peptide side chains. The issue of artifacts induced by short cutoff on the dynamical properties of hydrated peptides has been addressed by Schreiber and Steinhauser,[43] but their conclusions are difficult to transpose to the system investigated here. First, the peptide examined by these authors contained charged amino acids, namely, four L-lysines, supplemented by four counterions. Second, the effects of truncating long-range, electrostatic interactions were considered for only 120 ps of a molecular dynamics trajectory, over 400 times shorter than that reported in this paper. In contrast, Brooks *et al.*[44] reported only small changes in structural and thermodynamic properties of polar solutions when they compared smoothed cutoff to the full Coulomb potential without truncation. In general, there is at present no evidence that different cutoff schemes influence equilibrium and dynamic properties of nonpolar solutes over long time scales.

The SHAKE algorithm[45] was employed to constrain the bond lengths and bond angles of water and hexane, as well as the bond lengths between heavy atoms and hydrogen atoms of the peptide, to their equilibrium values, thereby removing high-frequency degrees of freedom. For the simulation in bulk water, the undecamer was placed in a cubic box with edge dimensions equal to 41.1 Å, containing 2162 water molecules. For the simulation in bulk hexane, the edge dimensions of the cubic box were equal to 45.5 Å. In this box, the peptide was surrounded by 394 hexane molecules. The potential energy functions used and the details of the simulations were the same as described above.

All the molecular dynamics simulations presented here were performed on one node of a four-processor Hewlett-Packard K-9000 (133 MHz) workstation at the NASA−Ames Research Center. For the study of the undecamer of poly-L-leucine in bulk water, the average CPU time per MD step was 0.629 s, making a total CPU time of 18.04 days for the 6.2 ns trajectory that was explored. In the case of the simulation in bulk hexane, the CPU cost per MD step was 0.252 s, which corresponds to a global CPU time of 4.37 and 5.39 days for 3.0 and 3.7 ns trajectories, respectively. Finally, the CPU investment for simulating the folding and the translocation of poly-L-leucine across the water−hexane interface was 0.577 s per MD step, hence making the total simulation time of 167.09 CPU days for the complete 50 ns trajectory.

3. Results

3.1. Folding and Unfolding of Poly-L-leucine in Bulk Water and Hexane. The undecamer of poly-L-leucine immersed in bulk water and in bulk hexane represents the end points in the transfer of the peptide across the water−hexane interface. On the basis of experimental data, concrete predictions can be made about the stability of the folded structures at these end points. This makes simulations of the end points especially valuable. If the conclusions from these simulations disagreed with the predictions, the results of all the subsequent efforts to describe the interfacial behavior of poly-L-leucine would not be reliable.

Two MD trajectories were obtained for poly-L-leucine in water. For one of them, the undecamer was placed in the aqueous medium as a β-strand. In this conformation, the carbonyl and amino groups of the backbone are readily accessible to the surrounding water, which should facilitate denaturation. As expected, the hydrated β-strand rapidly collapses into a random coil—*i.e.*, a family of disordered conformations observed over 6.2 ns of the MD trajectory. The final structure from this simulation was used in subsequent calculations whenever random coil was required as an initial configuration.

In the second trajectory, the initial structure of poly-L-leucine is α-helical. During *ca.* 10 ns of this trajectory, the peptide remains in approximately the same conformation. However, in separate simulations of the same system, aimed at determining α-helical propensity from the N-terminus,[46] some unfolding is observed at the C-terminus. These results imply that the helical structure of the undecamer is unstable in water, but only marginally. This is in line with experimental data on short, hydrophobic peptides. Although no experimental results are available on L-leucine-based peptides, largely due to their poor solubilities in water, it has been shown that alanine-based peptides form stable α-helices in water.[47,48] It was also determined from studies on synthetic helical dimers[11,49] and from host−guest peptide experiments[50] that L-leucine is a good helix-forming residue.

In contrast with the situation in water, poly-L-leucine, initially assigned a β-strand or a random coil conformation, refolds into a helical structure in bulk hexane within 3.0 or 3.7 ns, respectively. In both cases, folding proceeds sequentially from the C-terminus, *retro*, *i.e.*, $i \rightarrow i - 3$ and $i \rightarrow i - 4$, intramolecular hydrogen bonds forming between N−H groups on the C-terminus side and C═O groups on the N-terminus side.[51] As has been observed in X-ray elucidated proteins, propagation in the opposite direction is far less likely to occur.[52] The helical structure remains stable for over 10 ns of the MD trajectory. Probably the most comprehensive measures describ-

(36) Van Buuren, A. R.; Marrink, S. J.; Berendsen, H. J. C. *J. Phys. Chem.* **1993**, *97*, 9206.
(37) Pohorille, A.; Wilson, M. A. *J. Chem. Phys.* **1996**, *104*, 3760.
(38) Chipot, C.; Wilson, M. A.; Pohorille, A. *J. Phys. Chem. B* **1997**, *101*, 782.
(39) Wilson, M. A.; Pohorille, A. *J. Am. Chem. Soc.* **1996**, *118*, 6580.
(40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(41) Andrews, L.; Johnson, G. L. *J. Chem. Phys.* **1983**, *79*, 3670.
(42) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.
(43) Schreiber, H.; Steinhauser, O. *J. Mol. Biol.* **1992**, *228*, 909.
(44) Brooks, C. L.; III; Pettitt, B. M.; Karplus, M. *J. Chem. Phys.* **1985**, *83*, 5897.
(45) Ryckaert, J.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(46) Kollman, P.; Dixon, R.; Cornell, W.; Fox, T.; Chipot, C.; Pohorille, A. The development/application of a "minimalist" force field using a combination of ab initio and experimental data. In *Computer simulation of biomolecular systems: Theoretical and experimental applications*; Van Gunsteren, W. F., Weiner, P. K., Eds.; Escom: Leiden, The Netherlands, 1997; pp 83−96.
(47) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5286.
(48) Scholtz, J. M.; Marqusee, S.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Santoro, M.; Bolen, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 95.
(49) Lyu, P. C.; Liff, M. I.; Marky, L. A.; Kallenbach, N. R. *Science* **1990**, *250*, 669.
(50) Li, S. C.; Deber, C. M. *Struct. Biol.* **1994**, *1*, 368.
(51) Creighton, T. E. Conformation properties of polypeptide chains. *Proteins, structures and molecular properties*; Freeman and Company: New York, 1993; pp 171−199.
(52) Stickle, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226*, 1143.

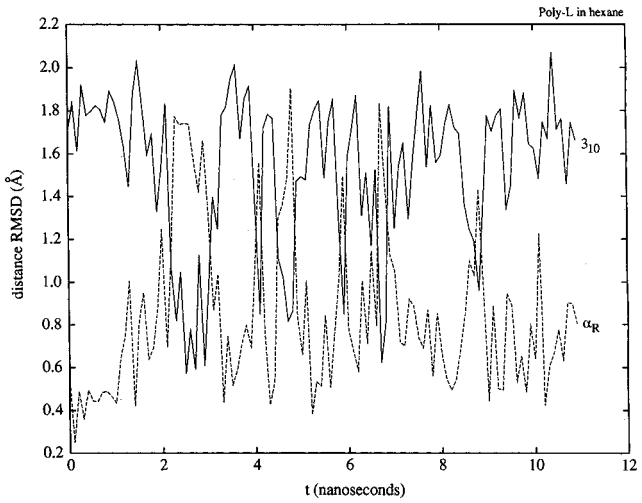

Figure 1. Distance RMSDs with respect to an ideal reference 3_{10} (solid line) and α (dashed line) helix[58] of the undecamer in bulk hexane as a function of time.

ing the folded structure are the distance root mean square deviations (RMSD) with respect to the ideal α ($\phi = -57°$, $\psi = -47°$) and 3_{10} ($\phi = -49°$, $\psi = -26°$) helices.[51] The two helices differ in the pattern of hydrogen bonding along the peptide backbone. In an α-helix, each hydrogen bond involves residues separated by three other residues, while in a 3_{10}-helix, the participating residues are separated by only two residues. Time evolutions of the two distance RMSDs are shown in Figure 1. Clearly, the peptide is primarily in the α-helical conformation, with occasional partial or complete conversions into the 3_{10}-helix. This is in agreement with experimental findings indicating that leucine-based and other hydrophobic peptides exist in organic solvents and in membrane bilayers as α-helices, distorted α-helices, or 3_{10}-helices.[6,7]

3.2. Peptide Folding and Translocation across the Interface. Several instantaneous configurations along the molecular dynamics trajectory describing the behavior of poly-L-leucine at the water−hexane interface are shown in Figure 2. As can be seen from this figure, the peptide translocates across the interface and undergoes folding into a helical structure.

One measure of peptide translocation across the interface is the position of its center of mass along the *z*-direction, perpendicular to the interface. Although the water−hexane interface undergoes some capillary wave fluctuations with time, it is locally sharp (almost no mixing occurs between water and hexane).[53] It is, therefore, possible to define its average location by considering the density profiles of water and hexane

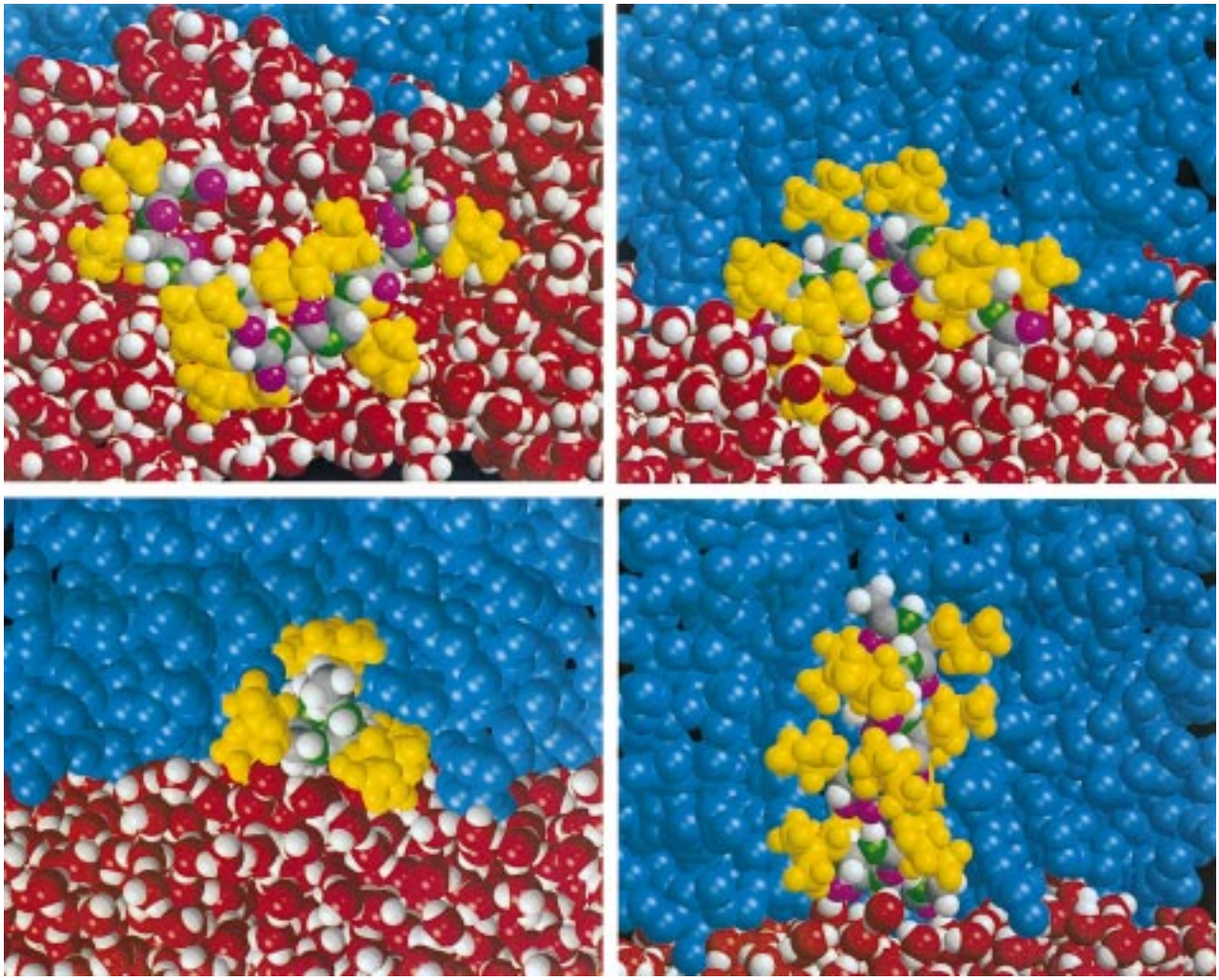

Figure 2. Four configurations of the undecamer of poly-L-leucine at the water−hexane interface: (a, top left) the initial, disordered structure on the water side of the interface; (b, top right) a turn after 21 ns; (c, bottom left) a 3_{10}-helix after 35 ns (view along the helical axis); (d, bottom right) an α-helix perpendicular to the interface. Oxygen atoms of water molecules are red, hydrogen atoms of water and the peptide backbone are white, methylene and methyl groups of hexane are blue, carbon, oxygen, and nitrogen atoms of the peptide backbone are gray, magenta, and green, respectively, and all atoms of the peptide side chains are yellow. The pictures were obtained using the program RASTER3D.[105,106]

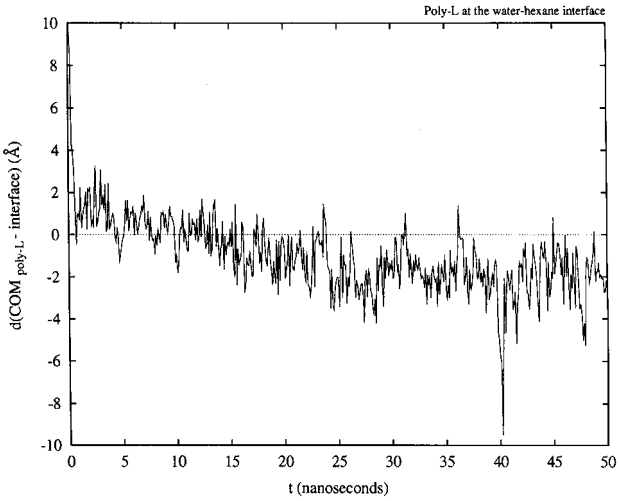

Figure 3. Position of the center of mass of the peptide along the *z*-direction as a function of time. The average position of the interface is at $z = 0$, and the water phase is at $z > 0$.

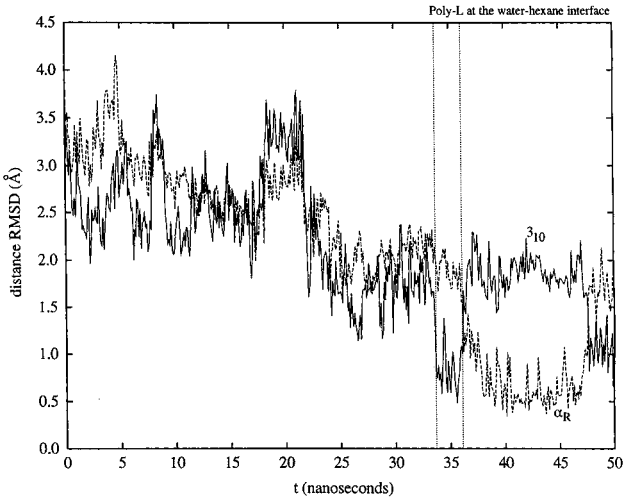

Figure 4. Distance RMSDs with respect to an ideal reference 3_{10} (solid line) and α (dashed line) helix[58] of the undecamer at the water−hexane interface as a function of time.

generated over different time periods along the trajectory. In Figure 3, it is shown how the position of the peptide changes as a function of time. Initially, the peptide in a random coil configuration, representing the final state in the simulations of the undecamer in water, as described above, was placed in water such that its center of mass was located 11 Å from the interface. However, as shown in Figure 2a, some atoms of the peptide were separated from the hexane phase by only 5 Å. This distance was set such that some hexane molecules were within the cutoff distance from these atoms, and therefore, poly-L-leucine still was still able to "see" the water−hexane interface.

In the first nanosecond of the MD trajectory, the peptide moves rapidly toward the interface. This rapid motion cannot be accounted for by simple diffusion but, instead, is probably better interpreted as diffusion in an external potential of mean force, which decreases toward the hexane phase. Once the peptide reaches the interface, its further movement along *z* becomes quite slow. The peptide largely translocates to hexane, but remains close to the water phase and does not desorb from the interface in the course of the simulation.

General features of interfacial folding of poly-L-leucine and its relation to peptide translocation across the interface are best characterized by distance RMSDs with respect to α- and 3_{10}-helices. As shown in Figure 4, their initial values are close to 3.5 Å, confirming that the starting structure of poly-L-leucine is quite disordered. In the first 34 ns of the MD trajectory, the distance RMSD is reduced by half, which indicates that the peptide undergoes partial folding. The reduction, however, is not monotonic; transient disordering is observed between 18 and 23 ns. By comparing Figures 3 and 4, a general trend is observed whereby the reduction in the distance RMSD is accompanied by peptide translocation across the interface. This suggests that folding and insertion into the nonpolar phase are closely coupled. At approximately 34 ns, the distance RMSD with respect to the 3_{10}-helix decreases to 0.5 Å, while the same quantity with respect to the α-helix remains unchanged. The situation reverses between 36 and 38 ns: the distance RMSD for the α-helix is reduced to 0.5 Å, and, simultaneously, that for the 3_{10}-helix increases to 1.7 Å. Toward the end of the simulation, another reversal takes place, whereby the distance RMSD increases for the α-helix and decreases for the 3_{10}-helix. The observed changes in the distance RMSD indicate that interfacial folding of poly-L-leucine to an α-helix proceeds through a short-lived intermediate, a 3_{10}-helix. This is shown in Figure 2c. Furthermore, the folded structure appears to be a mixture of α- and 3_{10}-helices rather than a single helix, as has been noted for short, alanine-based peptides, through NMR and double-spin-labeled ESR experiments.[54−57]

3.3. Detailed Mechanism of Interfacial Folding. To examine the mechanism by which poly-L-leucine folds during the computer simulation, several quantities are considered. Each of them probes one characteristic of the peptide as a function of time, but in combination, they provide a detailed description of the folding process. Changes in the ϕ and ψ angles for all 11 peptide units yield detailed information about conformational transitions in the peptide backbone. These conformational changes can be related to folding by monitoring the nine and ten distances between carbonyl oxygen and amide hydrogen atoms that form intramolecular hydrogen bonds in the α- and 3_{10}-helices, respectively. The evolution of these distances with time in the 3_{10}-helix is shown in Figure 5. When the distance is smaller than 2.3 Å, we consider the atoms involved to be hydrogen bonded.

The detailed characteristics of peptide folding are supplemented by two collective quantities—the macrodipole of the undecamer, μ, and the asymmetry coefficient, A_3. They are shown in Figures 6 and 7, respectively. Since each folded residue contributes *ca.* 3.2−3.5 D to the total molecular dipole moment of the peptide,[58,59] μ is a useful measure that relates folding to peptide polarity. A_3, in turn, provides global information about the shape of the peptide.[60] It is calculated from the three radii of gyration, R_0^k that are defined as

$$R_0^k = \sqrt{I^k/M} \qquad (1)$$

(53) Pohorille, A.; Wilson, M. A. *J. Mol. Struct.: THEOCHEM* **1993**, *284*, 271.
(54) Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L. *Nature* **1992**, *359*, 653.
(55) Fiori, W. R.; Miick, S. M.; Millhauser, G. L. *Biochemistry* **1993**, *32*, 11957.
(56) Millhauser, G. L. *Biochemistry* **1995**, *34*, 3873.
(57) Millhauser, C. J.; Hanson, P.; Bolin, K. A.; van de Ven, F. J. M. *J. Mol. Biol.* **1997**, *267*, 963.
(58) Creighton, T. E. *Proteins, Structures and Molecular Properties*; Freeman and Company: New York, 1993.
(59) Smythe, M. L.; Huston, S. E.; Marshall, G. R. *J. Am. Chem. Soc.* **1995**, *117*, 5445.
(60) Rudnick, J.; Gaspari, G. *Science* **1987**, *237*, 384.

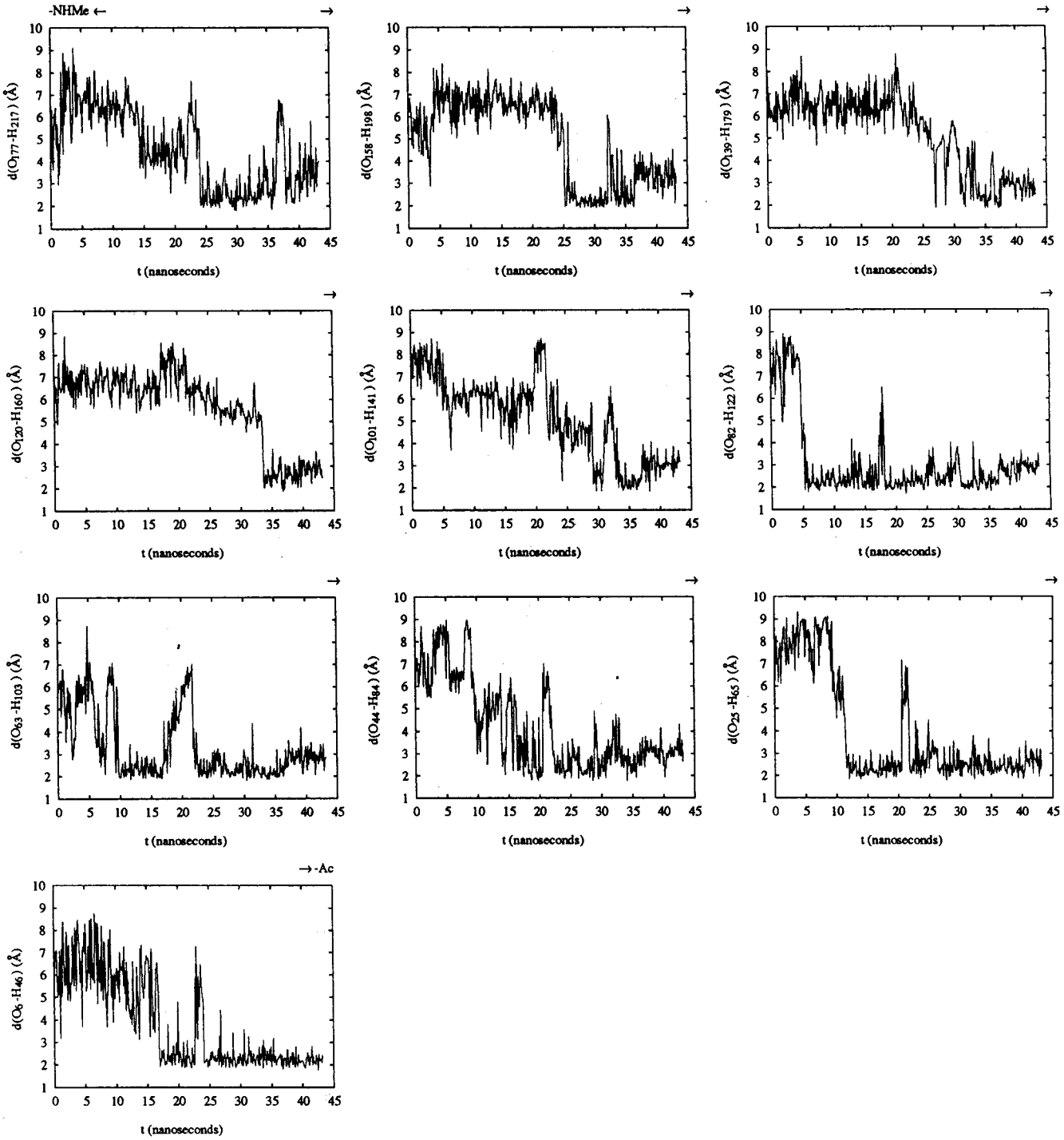

Figure 5. Time evolution of ten C═O···H−N distances that can form intramolecular hydrogen bonds in the 3_{10}-helix of the undecamer.

where $\{I^k\}$ are the eigenvalues of the poly-L-leucine inertia tensor and M is the total mass of the peptide. Once $\{R_0^k\}$ are known, A_3 is calculated as

$$A_3 = \frac{\sum_{i>j}^{3} \langle ((R_0^i)^2 - (R_0^j)^2)^2 \rangle}{2\langle (\sum_{i=1}^{3} (R_0^i)^2)^2 \rangle} \qquad (2)$$

A_3 takes values between 0 and 1, with 0 corresponding to the perfectly spherical shape and 1 to the linear shape.

The analysis of the quantities described above provides a detailed description of the folding process observed in our simulations. The first five intramolecular hydrogen bonds in the peptide backbone are formed at the N-terminus of the peptide between 6 and 17 ns of the MD trajectory. Their formation, however, does not proceed sequentially. All of the bonds are characteristic of the 3_{10}-helix. Accompanying changes in the total dipole moment of the peptide and in the asymmetry coefficient reflect the observed partial folding. The increase in μ by 15−20 D and the increased peptide linearity are consistent with the formation of a helical region spanning half of the undecamer.

Between 18 and 23 ns, the already formed hydrogen bonds in the backbone become intermittently broken and, subsequently, re-form. This brief disordering is accompanied by a transient increase in the distance RMSD with respect to the helical

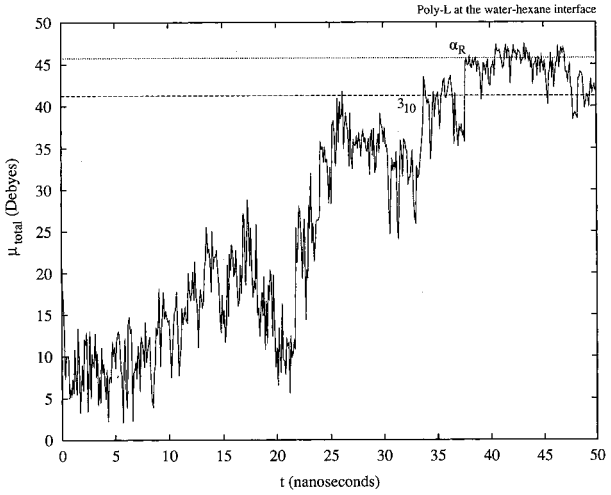

Figure 6. Time evolution of the total dipole moment of the peptide. Dotted and dashed lines correspond to the total dipole moment of the α- and 3_{10}-helices, respectively.

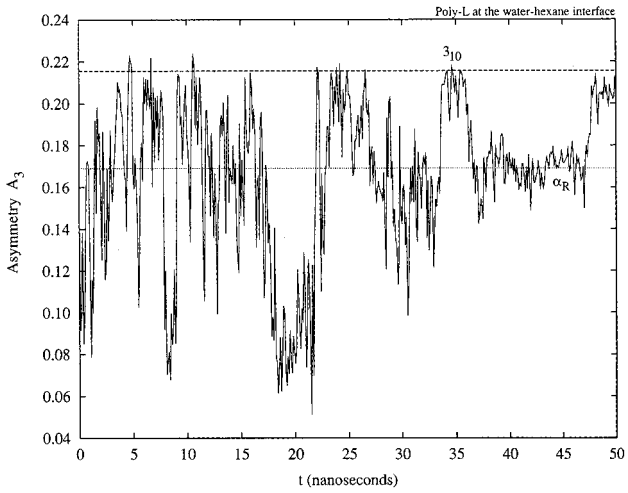

Figure 7. Time evolution of the asymmetry coefficient A_3 of the peptide. Dotted and dashed lines correspond to the total dipole moment of the α- and 3_{10}-helices, respectively.

structures and a reduction of the peptide macrodipole to the range of values characteristic of a random coil. The temporary interconversion to a disordered structure is caused by a transition in the backbone at the fifth residue from the N-terminus to the ϕ and ψ angles characteristic of the C_{7-ax} conformation. As shown in Figure 2b, the subsequent changes in the ϕ and ψ angles at the sixth residue create a turn in the peptide. As manifested by a large decrease in A_3, it yields a more spherical, compact structure. Once the ϕ and ψ angles at the fifth residue return to the values typical of helical structures, the N-terminal half of the peptide rapidly refolds to the 3_{10}-helix.

Further folding proceeds sequentially from the C-terminus. After 34 ns, μ and A_3 reach the values expected for fully helical conformations and their fluctuations diminish, which indicates that the folded peptide is considerably less flexible than disordered or partially ordered structures.

Before the MD trajectory is terminated at 50 ns, a transition from the 3_{10}- to the α-helix and a reversed transition from the α- to the 3_{10}-helix are observed. Both are approximately, although not precisely, sequential, but proceed from the opposite ends. The first transition is initiated from the N-terminus, whereas the second one starts from the C-terminus. Both μ and A_3 appear to be sufficiently sensitive to distinguish between the two types of helices and to identify when the interconversions occur. For the α-helix, μ is larger[59] and A_3 is smaller, compared to those for the 3_{10}-helix. During transitions between these two helices the values of both quantities drop below the values for either helix.

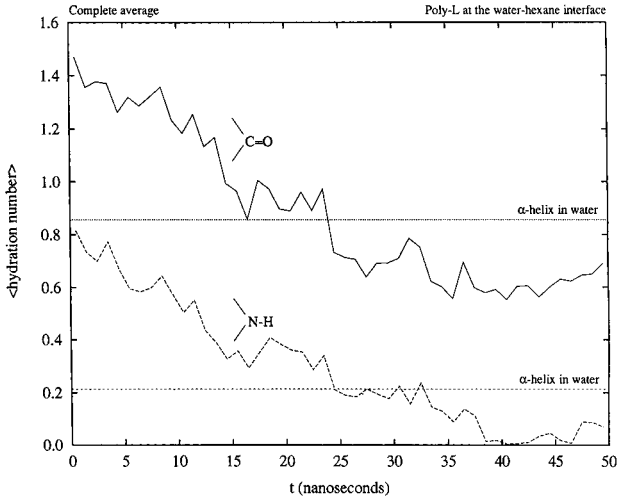

Figure 8. Changes in the hydration numbers along the molecular dynamics trajectory, averaged over all oxygen (solid line) and all nitrogen (dashed line) atoms of the peptide.

Both folding and translocation from an aqueous solution to hexane impose restrictions on the access of water to different groups in the peptide. Of particular interest is the hydration of the polar C═O and N─H groups in the peptide backbone because of their possible involvement in intramolecular hydrogen bonding stabilizing helical structures. Excluding aqueous solvent from direct contact with these groups favors hydrogen bonds along the backbone, by removing competition from hydrogen bonding donors and acceptors in water molecules. The degree of hydration can be conveniently measured as the hydration number, defined here as the average number of water molecules, the oxygen atoms of which are within, respectively, 3.25 and 2.95 Å of a carbonyl oxygen or an amide nitrogen atom in the peptide. The calculated hydration numbers are averaged over 1 ns intervals. Changes of the hydration numbers, averaged over all oxygen and nitrogen atoms, along the MD trajectory, are shown in Figure 8. Over the whole 50 ns, the oxygen atoms are better hydrated than nitrogen atoms by approximately 0.6 water molecule. In both cases, the initial hydration numbers are higher than the corresponding numbers calculated for the α-helical peptide in water by about 0.6. This is consistent with the expectations that water molecules access polar groups better in a random coil than in a helical structure. As the peptide folds and translocates across the interface, the hydration numbers decrease. For the folded peptide, the average hydration number of the oxygen atoms is reduced to 0.6. Whereas oxygen atoms remain partially exposed to water, nitrogen atoms, even those not involved in hydrogen bonding at the N-terminus of the peptide, are completely dehydrated. The hydration numbers for both oxygen and nitrogen atoms are approximately 0.2 less than the corresponding numbers for the α-helix in water, suggesting that the hexane phase provides some shielding of intramolecular hydrogen bonds in the interfacially located peptide.

The analysis of peptide hydration can be further refined by calculating hydration numbers separately for polar atoms participating and not participating in the hydrogen bonds at a given time. The results of these calculations for oxygen atoms

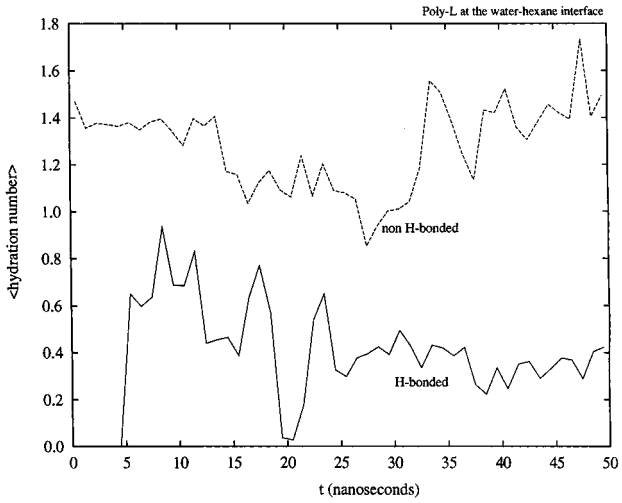

Figure 9. Hydration numbers for oxygen atoms of the peptide that participate (solid line) and do not participate (dashed line) in the hydrogen bonds, shown as functions of time.

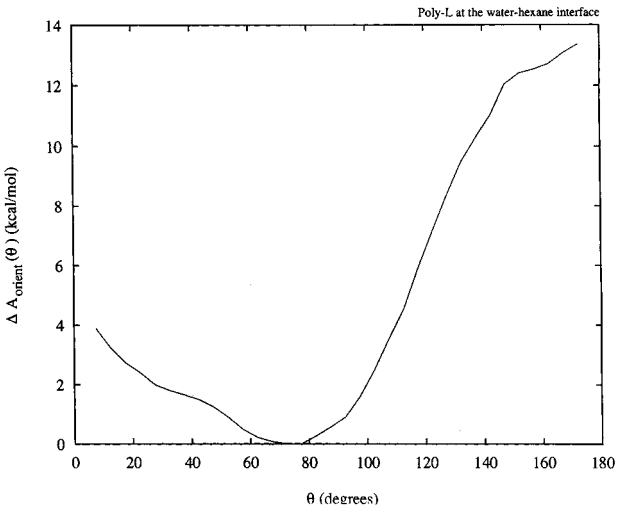

Figure 10. Orientational free energy profile of the helical peptide at the water−hexane interface. At 0° and 180°, the peptide is perpendicular to the interface with the C- and N-termini exposed to water, respectively. At 90° the peptide is parallel to the interface.

are shown in Figure 9. Note that since no hydrogen bonds exist during the first 5 ns and, again, at approximately 20 ns, the hydration numbers for the hydrogen-bonded oxygen atoms at these time intervals are zero. Also, fluctuations in the average hydration number for these atoms are large whenever the number of hydrogen bonds is small. Taking this into account, differences in the hydration numbers with time for both bonded and nonbonded oxygen atoms are found to be statistically insignificant. These hydration numbers, averaged over time, are 0.4 ± 0.15 and 1.3 ± 0.1, respectively. Thus, oxygen atoms participating in hydrogen bonding are always fairly well, albeit not completely, screened from water. Their hydration numbers are reduced, on average, by 0.3, compared to the hydration numbers for the α-helix in water. In contrast, hydration numbers for oxygen atoms not involved in hydrogen bonding are the same as for similar atoms in the random coil or the α-helix in an aqueous solution. This, again, suggests that the formation of intramolecular hydrogen bonds is closely coupled to the dehydration of the participating atoms.

3.4. Hydrophobic Helices at Aqueous Interfaces. To describe the interfacial behavior of the folded peptide, two questions are particularly relevant: (i) What is the free energy of the undecamer as a function of its orientation with respect to the interface? (ii) What is the free energy of transferring the peptide from the interface to the bulk hexane? The results from the 50 ns trajectory offer only partial answers to these questions. As has already been pointed out, the undecamer remains adsorbed at the interface. This implies that desorption into hexane either is strongly unfavorable or requires surmounting a high free energy barrier.

At the interface, the most probable orientation of the peptide is nearly parallel to the interface. The undecamer, however, exhibits considerable orientational flexibility. Most commonly, it rotates in such a way that the N-terminus is buried in hexane and, occasionally, even becomes perpendicular to the interface. One such case is shown in Figure 2d. When the peptide adopts this perpendicular orientation, its center of mass moves nearly 9 Å into the hexane phase while its C-terminus still remains located at the interface. This explains the large downward spike observed in Figure 3 around 40 ns. Whereas orientations in which the N-terminus is immersed in hexane are observed in the simulations, practically no information about the reversed orientations is available, suggesting that they may be unfavorable.

Since changes in both orientation and position with respect to the interface may be associated with large changes in the free energy, probing them in a single MD trajectory is inefficient. Instead, a different strategy must be employed. The free energy, $\Delta A_{\text{orient}}(\theta)$, as a function of the angle θ between the helical axis of the peptide and the normal to the interface was calculated using the "umbrella sampling" method.[61] For the peptide parallel to the interface, θ is equal to 90°, and for the peptide perpendicular to the interface, with the N-terminus buried in hexane, $\theta = 0°$. The full range of θ was divided into a sequence of "windows"[61,62] 40−50° wide. Two consecutive windows overlapped by at least 15°. Six windows were needed to obtain a complete free energy profile. For each window, a MD trajectory 1−2 ns long was obtained. The only exception was the range of 40−90°, in which the results from the last 10 ns of the 50 ns trajectory were used. A biasing potential $U_b(\theta)$, of the form $k_\theta|\theta - \theta_{\text{min}}|^2$ and $k_\theta'|\theta_{\text{max}} - \theta|^2$, where θ_{min} and θ_{max} are the bounds of the window and k_b and k_b' are force constants, was added in some windows to ensure a more uniform sampling of all orientations within these windows and, therefore, improve the statistical accuracy of the results. In each window, $\Delta A_{\text{orient}}(\theta)$ was calculated from the probability $P(\theta)$ of finding the center of mass of the solute at an angle θ within the window:

$$\Delta A_{\text{orient}}(\theta) = -k_B T \ln P(\theta) + U_b(\theta) \qquad (3)$$

where k_B is the Boltzmann constant and T is the temperature. The free energy profile in the full range of θ was obtained by exploiting the requirement that $\Delta A_{\text{orient}}(\theta)$ must be a continuous function of θ.

In all windows the peptide remains helical and adsorbed at the interface. The calculated free energy profile is shown in Figure 10. In agreement with the observations from the 50 ns trajectory, the minimum in $\Delta A_{\text{orient}}(\theta)$ corresponds to the peptide orientation nearly parallel to the interface. The most striking feature of the profile is its asymmetry. Rotating the peptide into the perpendicular orientation, in which the N-terminus is buried in hexane, requires only 4 kcal/mol. In contrast, rotating the peptide in the opposite direction requires 13 kcal/mol. Thus, the insertion of the undecamer into a nonpolar phase from the C-terminus should be highly unlikely.

(61) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187.
(62) Valleau, J. P.; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457.

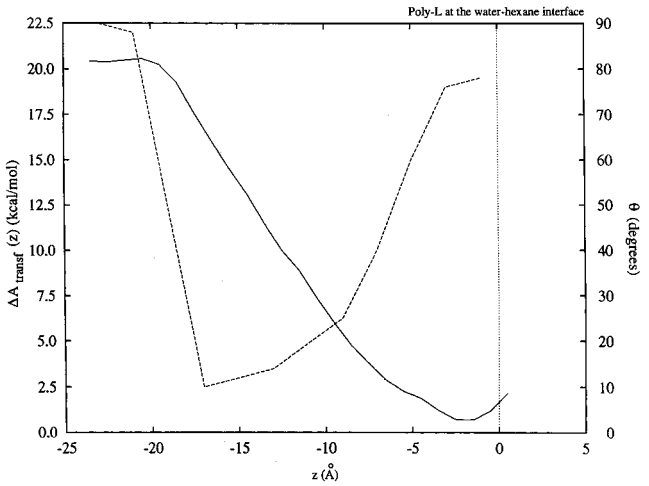

Figure 11. Free energy profile of transfer of the peptide from the water−hexane interface to the middle of the hexane lamella (solid line) and the corresponding changes in the average value of θ (dashed line) The average position of the interface is at $z = 0$, and the center of the hexane lamella is at −25 Å.

The free energy of transferring the peptide from the interface to hexane, $\Delta A_{transf}(z)$, along the z-direction was calculated using the same approach. Again, the complete simulation was divided into six windows. Each window was 5−6 Å wide, and two consecutive windows overlapped by 2 Å. In every window not only the probability of finding the center of mass of the peptide at position z was calculated, but also the orientational distribution was monitored.

Both $\Delta A_{transf}(z)$ and the average value of θ are shown in Figure 11. The free energy profile exhibits an interfacial minimum and increases monotonically from the interface to the position approximately 5 Å from the center of the hexane lamella, which indicates that desorption is not associated with a free energy barrier. Near the center of the hexane lamella $\Delta A_{transf}(z)$ is constant, as expected in a bulk medium. The free energy of transfer from the interface to the center of the lamella is substantial and amounts to almost 20 kcal/mol.

The changes in orientational preferences of the undecamer as a function of its position relative to the interface shed light on the mechanism by which the peptide inserts into a nonpolar phase. As the center of mass of the undecamer is moved into hexane, the preferred peptide orientation shifts from parallel to perpendicular to the interface. Simultaneously, the distribution of θ becomes progressively sharper, thereby indicating that the peptide loses most of its rotational flexibility. Only when the center of mass is about 5 Å from the center of the lamella does the undecamer rapidly regain its rotational freedom and the distribution of θ characteristic of a molecule in a bulk solvent. The region of transition from the rigid to random orientation coincides with the range of z values, in which $\Delta A_{transf}(z)$ becomes flat.

The observed changes in orientational preferences of the peptide can be explained by considering its pattern of hydration. As has already been mentioned, the oxygen atoms of the backbone, which do not participate in intramolecular hydrogen bonds, remain well hydrated at the interface, even in the fully folded conformation. These atoms are located at the C-terminus of the peptide. As the center of mass of the peptide is shifted into hexane, the atoms can retain partial hydration only if the peptide swings from the parallel to perpendicular orientation. The deeper the center of mass is buried into hexane, the more rigid the peptide must be to expose its C-terminus to water. Only when the peptide is moved into hexane sufficiently deeply so that the C-terminus does not reach the interface—even in a perfectly perpendicular alignment—does it fully regain rotational freedom.

4. Discussion

4.1. Peptide Folding and Translocation. Even though the water−hexane interface is considerably simpler than water−membrane systems, the results of the present simulations bear a clear relationship to several central ideas about the folding and insertion into bilayers of hydrophobic peptides or protein fragments. At the same time, these simulations reveal additional characteristics of these processes that have not yet been adequately studied. In particular, they point out the essential role of the interface in mediating folding of peptides that are disordered in aqueous solution.

The significance of the interfacial environment is illuminated by the finding that the undecamer of poly-L-leucine is interfacially active, *i.e.*, exhibits a free energy minimum at the interface, even though it is entirely built of hydrophobic residues. This phenomenon of interfacial activity, well-documented for amphiphilic peptides,[10,63,64] has also been observed experimentally in water−bilayer systems for a broad range of short peptides that do not have well-separated hydrophobic and hydrophilic sides.[15,65] Similar results were obtained in computer simulations of terminally blocked amino acids at the water−hexane[66,67] and water−membrane[68] interfaces, blocked dipeptides at the water−hexane interface,[69] and tripeptides at the water−membrane interface.[70] In fact, interfacial activity is characteristic not only of small peptides but also, in general, of many solutes that have both polar and nonpolar regions, as demonstrated in several recent experimental[71−73] and computer simulation[38,74] studies. The tendency to accumulate at the interface, exhibited by all these solutes, results mostly from a balance between two oppositely changing contributions to the free energy.[37] The reversible work needed to create a cavity that can accommodate the solute is lower in the nonpolar phase than in water whereas the contribution from electrostatic solute−solvent interactions is lower in water. A similar picture emerges from the recent computational study of the transfer of the α-helical conformation of poly-L-alanine from an aqueous solution into a lipid bilayer.[23] Using a dielectric continuum representation of both water and the membrane, a deep free energy minimum is found for the horizontal adsorption of a

(63) Blondelle, S. E.; Ostreh, J. M.; Houghten, R. A.; Perez-Paya, E. *Biophys. J.* **1995**, *68*, 351.
(64) Cornut, I.; Desbat, B.; Turlet, J. M.; Dufourcq, J. *Biophys. J.* **1996**, *70* 305.
(65) Brown, J. W.; Huestis, W. H. *J. Phys. Chem.* **1993**, *97*, 2967.
(66) Pohorille, A.; Wilson, M. A. Reaction dynamics in clusters and condensed phases. In *The Jerusalem symposia on quantum chemistry and biochemistry*; Jortner, J., Levine, R. D., Pullman, B., Eds.; Kluwer: Dordrecht, 1993; Vol. 26, p 207.
(67) Chipot, C.; Pohorille, A. *J. Phys. Chem. B* **1998**, *102*, 281.
(68) Pohorille, A.; Wilson, M. A. Interaction of a model peptide with a water-bilayer system. In *Structure and reactivity in aqueous solution: Characterization of chemical and biological systems*; Cramer, C. J., Truhlar, D. G., Eds.; ACS Symposium Series No. 568; ACS: Washington, DC, 1994; Chapter 11, p 395.
(69) Chipot, C.; Pohorille, A. *J. Mol. Struct.: THEOCHEM* **1997**, *398*/*399*, 529.
(70) Damodaran, K. V.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1995**, *117*, 6561.
(71) Baber, J.; Ellena, J. F.; Cafiso, D. S. *Biochemistry* **1995**, *34*, 6533.
(72) North, C.; Cafiso, D. S. *Biophys. J.* **1997**, *72*, 1754.
(73) Xu, Y.; Teng, P. *Biophys. Biochim. Acta* **1997**, *1323*, 154.
(74) Pohorille, A.; Wilson, M. A.; Chipot, C. *Prog. Colloid Polym. Sci.* **1997**, *103*, 29.

helix at the interface. This helix position corresponds to the most favorable balance between the hydrophobic effect, which provided the major driving force for helix insertion, and the large free energy penalty associated with dehydration of peptide hydrogen bonds, which opposes the insertion. For highly flexible solutes this is only a qualitative, simplified picture. Other contributions to the free energy, due to conformational transitions and immobilization at the interface also influence interfacial activity of solutes.[23] In water−bilayer systems, additional effects, arising from the collective motions associated with membrane reorganization, as well as transmembrane potentials, should also be considered.[30,75,76]

In the most commonly accepted models of folding transmembrane proteins,[15,77,78] it has been postulated that once the peptide becomes adsorbed at the interface it undergoes folding, followed by insertion into the bilayer. A similar picture emerges from simulations of membrane insertion of simple lattice protein chains.[22] The present calculations are, in general, consistent with this picture. They demonstrate that folding takes place in the interfacial environment rather than in the nonpolar phase. They, however, also indicate that some penetration into hexane is required for folding and that these two processes are tightly coupled. Polar groups exposed to hexane become partially dehydrated and can form stable, intramolecular hydrogen bonds along the backbone. Accompanying local folding, the nearby, hydrophobic side chains show a clear tendency to shield the polar groups of the backbone from the solvent. This, in turn, promotes further penetration of the undecamer into the nonpolar phase. When folding is completed, the peptide is mostly immersed in hexane, but still retains some direct interactions with water.

Recently, an alternative model has been proposed to explain the results of fluorescence and stopped-flow circular dichroism measurements of membrane insertion of a hydrophobic signal peptide.[79] In this model, the peptide does not adopt an α-helical structure upon binding to the membrane, but only after it becomes inserted into the bilayer. Since the conclusions are based on a difficult interpretation of experimental data, more studies are warranted on this peptide. More importantly, perhaps, the model fails to explain the energetics of transferring polar groups in the backbone, strongly hydrated in the disordered state, into the nonpolar interior of the membrane.

The detailed mechanism of peptide folding observed in our simulations is at variance with the standard theories of helix−coil transitions of peptides in bulk media.[80,81] In these theories, helix formation is divided into two stages: the helix initiation and propagation. Once the initial intramolecular hydrogen bond is formed in the first stage, further helix growth is cooperative and sequential. This view has sound experimental bases arising from kinetic measurements of helix−coil transition in aqueous solutions.[82−84] The simulated, interfacial folding of poly-L-leucine also exhibits some cooperativity but, in contrast with the standard models, does not proceed sequentially and does not start from the end of the peptide. This surprising result does not seem to be an artifact of inadequate potential energy functions describing peptide conformations, since folding of the undecamer in hexane proceeds sequentially from the C-terminus side. This is in agreement with an apparent tendency of protein helices to fold or unfold from the C-terminus. This tendency was observed experimentally for the H-helix of myoglobin[85] and for the C-helix of myohemerythrin.[86] The same tendency was also noted in MD simulations of the S-peptide of ribonuclease A in aqueous solution[87] and the helix propagation in alanine peptides.[27] Similarly, as has been found in this work, unfolding of poly-L-leucine in water appears to start from the C-terminus.

In view of the similarities between the mechanism of folding in bulk media found in this work and in previous studies, it is natural to propose that the nonsequential folding of poly-L-leucine is induced by the interfacial environment. A possible explanation of this phenomenon relies on the idea that shielding the polar groups of the backbone from water promotes the formation of intramolecular hydrogen bonds. According to this hypothesis, it is expected that dehydration of the C═O and N−H groups, which does not have to occur initially at the ends of the peptide, precedes local folding. The hypothesis can be tested by appealing to hydration patterns around polar groups. If dehydration of the donor and the acceptor is, indeed, a prerequisite for the formation of a stable hydrogen bond at the interface, then the average hydration numbers around these atoms in a time interval (*e.g.*, 1 ns long) shortly before they form a hydrogen bond should be markedly lower than the average hydration numbers around atoms of the same type not involved in hydrogen bonding. For carbonyl oxygen atoms, these two numbers are equal to 0.65 ± 0.2 and 1.3 ± 0.1, respectively. For comparison, the average hydration number for the oxygen atoms engaged in hydrogen bonding is 0.4 ± 0.15. This comparison shows that, indeed, dehydration precedes, rather than results from, the formation of hydrogen bonds.

Another interesting feature of interfacial folding observed in the present set of simulations is its highly dynamic nature. It appears that helical fragments can fold and unfold before the full-length helix is formed. One such case is observed at the N-terminus of the peptide, when five consecutive hydrogen bonds are briefly broken to form a turn and then re-form approximately 5 ns later. This dynamic behavior can be interpreted in terms of the nascent helix, an ensemble of transient, partially ordered secondary structures, which rapidly interconvert by way of unfolded states. The concept of the nascent helix was introduced to explain conformational equilibria in the C-helix of myohemerythrin in water and organic solvents.[86] In general, the unfolded or partially folded peptide exhibits considerable flexibility, as can be seen, for example, in Figure 7. In contrast, the folded undecamer appears to be rather rigid, although the C-terminal part, but not the N-terminal part, shows some degree of disorder, in line with the similar behavior of other peptides in water.[25,85,87] The ends are only weakly frayed, even though L-leucine side chains cannot form hydrogen bonds with the backbone and, therefore, cannot be considered as good capping residues.[88,89] It was, however,

(75) Beschiaschvili, G.; Seelig, J. *Biochemistry* **1992**, *31*, 12813.
(76) Wimley, W. C.; White, S. H. *Biochemistry* **1993**, *32*, 6307.
(77) Popot, J. L.; Gerchman, S. E.; Engelman, D. M. *J. Mol. Biol.* **1987**, *198*, 655.
(78) Lemmon, M. A.; Engelman, D. M. *Q. Rev. Biophys.* **1987**, *27*, 157.
(79) Golding, C.; Senior, S.; Wilson, M. T.; O'Shea, P. *Biochemistry* **1996**, *35*, 10931.
(80) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526.
(81) Poland, D.; Scheraga, H. A. *Theory of helix-coil transitions*; Academic Press: New York, 1970.
(82) Altmann, K. H.; Wojcik, J.; Vasquez, M.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 107.
(83) Chakrabartty, A.; Schellman, J. A.; Baldwin, R. L. *Nature* **1991**, *351*, 586.
(84) Scholtz, J. M. Baldwin, R. L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95.
(85) Soman, K. V.; Karimi, A.; Case, D. A. *Biopolymers* **1991**, *31*, 1351.
(86) Dyson, H. J.; Rance, M.; Houghton, R. A.; Wright, P. E.; Lerner, R. A. *J. Mol. Biol.* **1988**, *201*, 201.
(87) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1991**, *30*, 3864.
(88) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648.
(89) Chakrabartty, A.; Doig, J. D.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11332.

pointed out that transmembrane helices may not require capping for good stability.[90]

Since only one folding event has been simulated, it is difficult to assess whether the highly dynamic, nonsequential mechanism of interfacial folding of hydrophobic peptides is predominant, or even typical. In fact, on the basis of the hypothesis about the role of local dehydration in promoting the formation of hydrogen bonds, it may be suggested that there is no unique folding pathway at the interface.

Similarly, a single event, in which the undecamer folds in 36 ns, is insufficient to determine the average, simulated folding time of this peptide at the water−hexane interface. It is, nevertheless, reasonable to expect that this time is on the order of 10−100 ns. A very similar time scale was recently recorded by Williams *et al.*,[91] in laser-induced temperature jump experiments, on a 21-residue alanine-based peptide in water. The time constant for unfolding was found to be 160 ± 60 ns, and the relaxation time for folding was approximated to be 16 ns. Computer simulations of alanine-rich peptides in water, which represented solvent effect only in an average fashion, also yielded short folding times, approximately 3−4 ns.[28,92] The rapid helix−coil interconversion implies that elements of secondary structure form before slower, long-range tertiary contacts are made. This view is consistent with the proposed mechanisms of protein insertion into membranes[15,16] and the two-stage model of folding transmembrane proteins, in which it is assumed that individual helices form first and then assemble into native protein.[9]

Besides the lack of statistics, there are other limits on our ability to interpret the observed folding time. First, potential energy functions employed in MD simulations, generally parametrized using quantum mechanical *ab initio* calculations of terminally blocked single amino acids, may not be equally accurate for larger systems. The recent LMP2/cc-pVTZ(-f) calculations of Beachy *et al.*[93] suggest that these potential energy functions are only moderately successful in reproducing conformational energies of the blocked L-alanine tripeptide. To reproduce these *ab initio* results, the original AMBER force field of Cornell *et al.*[40] was recalibrated such that the good agreement with quantum mechanical results found previously for blocked L-alanine single amino acid was retained.[46] Employing both the standard and the revised potential energy functions, the free energy for unfolding the first residue of the α-helical undecamer of poly-L-leucine from its N-terminus end in an aqueous solution was computed. Although the standard force field of Cornell *et al.*, employed in this work, reproduces the results of Beachy *et al.* better than most of the other force fields,[93] the two unfolding simulations indicate that it may have a tendency to overestimate the stability of α-helices.

The second limitation in interpreting the folding time is associated with uncertain accuracy of time scales for helix−coil transitions predicted from computer simulations. This is because potential energy functions for peptides were developed to describe correctly locally stable structures, but not necessarily the energetics of conformational transitions. Third, time scales of folding at the water−hexane interface may not be directly related to time scales at the water−membrane interfaces, due to the absence of slowly relaxing, collective motions of lipid molecules accompanying conformational and orientational changes in peptides. Finally, time scales of interfacial folding may depend strongly on peptide sequence because of the possibilities for strong electrostatic interactions between hydrophilic amino acids in disordered structures and lipid head groups or water. These interactions may be responsible for unusually slow α-helix formation in a part of bacteriorhodopsin, which is in apparent violation of the two-stage model.[94]

4.2. Roles of α- and 3_{10}-Helices. The present simulations illuminate an important role of the 3_{10}-helix in peptide folding, both as a folding intermediate and as a participant in the ensemble of folded structures. This role has been recently probed in several experimental and computational studies. In double-label electron spin resonance experiments on short trilysine-substituted poly-L-alanine peptides in water,[47] Millhauser and co-workers[54,55] found regions in 3_{10}-helical conformation. From these results, Millhauser[56] inferred that 3_{10}-helices constituted an important intermediate in the folding pathway from a random coil to an α-helix:

$$\text{random coil} \leftrightarrow \text{nascent helix} \leftrightarrow 3_{10}\text{-helix} \leftrightarrow \alpha\text{-helix} \quad (4)$$

This scheme is consistent with the previous work of Sundaralingam and Sekharudu,[95] who suggested, on the basis of their analysis of protein crystal structures, that 3_{10}-helices were trapped intermediates along folding and unfolding pathways. Computer simulations of peptide unfolding led to similar conclusions.[24,26,85] It was also proposed that transitions between α- and 3_{10}-helices play a role in enzyme dynamics.[96,97]

The relative equilibrium stabilities of α- and 3_{10}-helical conformations of short peptides have not been unequivocally established, and the balance between different factors that influence these stabilities is not fully understood. Recent NMR studies on alanine-rich peptides in water reveal a considerable population of the 3_{10}-helix, which is approximately equal to 25% in the middle of the peptide and reaches 50% at its ends.[57] In contrast, Smythe *et al.*[59,98] predict that the α-helix is more stable than the 3_{10}-helix in both polar and nonpolar environments. They proposed that the α-helix is energetically stabilized, primarily by solute−solvent interactions, and the 3_{10}-helix is preferred entropically. Consequently, water favors the α-helix, whereas nonpolar environments, such as membranes and the interior of proteins, could provide stabilizing conditions for the 3_{10}-helix. The free energy difference between the 3_{10}- and the α-helical conformations of the Ac− and NHMe− terminally blocked decamer of α-methylalanine was estimated at 7.6 kcal/mol in water, but only 3.2 kcal/mol *in vacuo*.[59,98] In agreement with these estimates, it was found in electron spin resonance investigations that alanine-based peptides are α-helical in water.[99] A similar conclusion follows from recent molecular dynamics simulations of the helix propagation in capped poly-L-alanine peptides, from both the N- and the C-terminus ends.[27] The absence of a minimum corresponding to the 3_{10}-helical state on the free energy surfaces in aqueous solution led to the suggestion that 3_{10}-helices may be kinetic, rather than thermodynamic, intermediates along the folding pathway.

(90) Presta, L. G.; Rose, G. D. *Science* **1988**, *240*, 1632.
(91) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. *Biochemistry* **1996**, *35*, 691.
(92) Kohn, W. D.; Mant, C. T.; Hodges, R. S. *J. Biol. Chem.* **1997**, *272*, 2583.
(93) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908.
(94) Riley, M. L.; Wallace, B. A.; Flitsch, S. L.; Booth, P. J. *Biochemistry* **1997**, *36*, 192.
(95) Sundaralingam, M.; Sekharadu, Y. C. *Science* **1989**, *244*, 1333.
(96) Gerstein, M.; Chothia, C. *J. Mol. Biol.* **1991**, *220*, 133.
(97) McPhalen, C. A.; Vincent, M. G.; Picot, D.; Jansonius, J. N.; Lesk, A. M.; Chothia, C. *J. Mol. Biol.* **1992**, *227*, 197.
(98) Smythe, M. L.; Huston, S. E.; Marshall, G. R. *J. Am. Chem. Soc.* **1993**, *115*, 11594.
(99) Smythe, M. L.; Nakaie, C. R.; Marshall, G. R. *J. Am. Chem. Soc.* **1995**, *117*, 10555.

Several studies indicate the presence of 3_{10}-helices in nonpolar media.[100−102] Perhaps the most relevant to the present work is the study of a peptide, composed of the alternating sequence of L-leucine and -alanine residues, in organic solvents and phospholipid bilayers.[6,7] The Fourier transform infrared data suggest that the peptide exists as a mixture of α- and 3_{10}-helical structures. This finding, as well as other results on the contribution of 3_{10}-helices to folding pathways and equilibrium populations of folded structures, is consistent with the role of these helices in interfacial folding, as identified in this work. It should, however, be borne in mind that the stability of the 3_{10}-helix may depend subtly on the peptide sequence and the exact nature of the environment.

The present simulation not only reveals the presence of the 3_{10}-helix in equilibrium with the α-helix, but also provides information about the mechanism of refolding between these two structures. In contrast with the interfacial formation of a 3_{10}-helix, this process is sequential in both directions. The transition from a 3_{10}- to an α-helix proceeds from the N-terminus, whereas the reverse process originates at the C-terminus. This result is consistent with MD simulations of L-alanine peptides in water, which indicate that the propensity to form a 3_{10}-helix is higher at the C-terminus than at the N-terminus.[27]

4.3. Insertion of the Folded Peptide into the Nonpolar Phase. Experimental studies indicate that hydrophobic peptides are most often incorporated into lipid bilayers as helices, in the transmembrane orientation.[6−8,103] The completely vertical insertion, in which a helix end is exposed to water at each side of the bilayer, is stabilized by hydration of the terminal carbonyl and amino groups which do not participate in intramolecular hydrogen bonding. Some peptides that are too short to span the membrane and, therefore, have only one terminus region exposed to water also tend to adopt perpendicular orientations.[103] Although transmembrane locations are prevalent, orientations parallel to the interface are also occasionally detected.[8,104] Similar conclusions follow from computer simulations. In MD simulations of a 32-residue poly-L-alanine in a fully hydrated dimyristoyphosphatidylcholine (DMPC) bilayer, it was found that the peptide remained essentially α-helical in the membrane interior, and was, on average, tilted 30° from the normal to the bilayer plane.[31] Transmembrane orientations of folded structures were also found to be favorable in lattice model calculations of peptide insertion into bilayers.[22] Calculations of the free energy of α-helical poly-L-alanine, parallel and perpendicular to the interface, led to the conclusion that both orientations are equally probable to within the accuracy of the approach.[23] A conceptually similar, but more approximate, method yielded the result that the transmembrane orientation of the 20-residue poly-L-alanine is more stable than the parallel orientation by 5 kcal/mol.[30]

Considering the differences between the water−hexane and the water−membrane interfaces, the results on orientational preferences of folded peptides are not directly transferable between these two systems. There are, nevertheless, clear similarities. Even though folded poly-L-leucine is preferentially located parallel to the water−hexane interface, the perpendicular orientation is also probable and, in fact, was observed during the simulations (see Figure 2d). The free energy difference between these two orientations is sufficiently small, so that the results of this work can be considered as being in general agreement with studies on peptides interacting with membranes. Furthermore, they clearly suggest a mechanism for peptide insertion into the membrane. The folded peptide, located in the interfacial region in the parallel orientation, incorporates into the bilayer by rotating into the perpendicular orientation. A very similar mechanism was proposed in previous experimental[15] and theoretical[21−23] studies.

A large free energy difference between the undecamer located perpendicularly to the interface with the N- and C-terminus exposed to water leads to a prediction that hydrophobic peptides penetrate the bilayer from the N-terminus. This tendency was also observed experimentally.[103] It should also be noted that hydrophobic signal sequences in proteins are typically located at the N-terminus.[3] The directionality of the insertion can be ascribed to a difference in the hydration of the carbonyl and the amino groups. As can be seen in Figure 8, the carbonyl groups that are not engaged in hydrogen bonding are markedly better hydrated than the corresponding amino groups. In the α-helix, there are three such C═O groups at the C-terminus and three N−H groups at the N-terminus. Thus, peptide insertion from the C-terminus requires a more extensive, energetically unfavorable dehydration of the polar groups than the insertion from the N-terminus. This further leads to the suggestion that the mode of insertion can be modulated by capping the C-terminus. The significance of unsatisfied hydrogen bonding capabilities at the ends of helices for the energetics of insertion was also pointed out by Ben-Tal *et al.*[23]

Peptide insertion into membranes should be clearly distinguished from its transfer from the interface into the nonpolar phase. The latter process requires a complete dehydration of the peptide and, therefore, is highly unfavorable. For the undecamer the free energy change in this process is nearly 20 kcal/mol. A similarly large free energy, between 30 and 35 kcal/mol, can be deduced for the transfer of a 25-residue poly-L-alanine helix from the interface to the interior of a nonpolar phase with a dielectric constant of 2.[23]

5. Conclusion

In the molecular dynamics simulations presented in this work, folding of the undecamer of poly-L-leucine at the water−hexane interface is studied at full atomic detail, without biases introduced by enforcing geometrical constraints on the peptide to define *a priori* the folding reaction coordinate, or increasing artificially the temperature of the system. Perhaps the major weakness of this work is the absence of additional simulations that would lend support to the observations made hitherto. Investigating, however, the folding and the translocation of nonpolar peptides across an aqueous interface requires a substantial CPU investment to analyze these phenomena over physically meaningful time scales. This explains why such simulations cannot be repeated in a routine fashion for a sufficiently large number of trajectories to form an appropriate statistical ensemble. Yet, although a single molecular dynamics trajectory is certainly insufficient to draw unequivocal, quantitive conclusions about folding time scales, it is still possible to predict general, qualitative trends for peptide folding at an interface. In particular, the results reported here remain in agreement with previous experimental and computational stud-

(100) Otoda, K.; Kitagawa, Y.; Kimura, S.; Imanishi, Y. *Biopolymers* **1983**, *33*, 1337.
(101) Basu, G.; Kuki, A. *Biopolymers* **1993**, *33*, 995.
(102) Augspurger, J. D.; Bindra, V. A.; Scheraga, H. A.; Kuki, A. *Biochemistry* **1995**, *34*, 2566.
(103) Otoda, K.; Kimura, S.; Imanishi, Y. *Biochim. Biophys. Acta* **1993**, *1150*, 1.
(104) Wu, Y.; He, K.; Ludtke, S. J.; Huang, H. W. *Biophys. J.* **1995**, *68*, 2361.
(105) Bacon, D. J.; Anderson, W. F. *J. Mol. Graphics* **1988**, *6*, 219.
(106) Merritt, E. A.; Murphy, M. E. P. *Acta Crystallogr.* **1994**, *D50*, 869.

ies. Simultaneously, they offer new insights into the mechanism by which hydrophobic peptides fold and insert into membranes. Folding is an interfacial phenomenon, and neither precedes nor follows peptide translocation from water to the nonpolar phase. Instead, these two processes are closely coupled. The nonpolar environment shields the carbonyl and the amino groups of the peptide backbone from water, thereby stabilizing intramolecular hydrogen bonding. The formation of such bonds, characteristic of folded structures, promotes, in turn, further partitioning of the peptide into the nonpolar phase. In contrast to folding in bulk hexane, the formation of helical structures at the interface is nonsequential and dynamic, with occasional, transient breaking of the already formed tracts of hydrogen bonds along the backbone. On the basis of these findings, it is hypothesized that there may be no unique pathway for interfacial folding.

Folding of poly-L-leucine at the interface does not proceed directly to the α-helix but, instead, involves a short-lived intermediate, the 3_{10}-helix. Occasional formation of this helix is also observed for the folded peptide. This suggests that the 3_{10}-helix not only mediates folding, but remains in an equilibrium with the α-helix once this process is completed. A similar conclusion was reached from studies of alanine-based peptides in an aqueous solution.[54−57]

The folded peptide is largely immersed in hexane, but remains in the interfacial region. This indicates that not only amphiphilic, but also hydrophobic, peptides, like the one investigated here, could be interfacially active. Although the preferred arrangement of the helical poly-L-leucine is parallel to the interface, the perpendicular orientation with the N-terminus buried in hexane is only slightly less favorable. This result lends support to a hypothesis that peptides assembled into helical structures at the water−membrane interface insert into the bilayer by rotating from a parallel to a transmembrane, perpendicular orientation.[15] In addition, it suggests that insertion proceeds from the N-terminus side, while the C-terminus remains at the interface. This is attributed to the high free energy cost for dehydrating the C-terminus carbonyl groups that are not involved in intramolecular hydrogen bonding.

Several results of this work may be relevant not only to the formation of α-helices at membrane surfaces, but also to the folding of water-soluble proteins. The environment in which the main elements of secondary structure are formed may be better approximated by the interface between water and a partially disordered, largely hydrophobic core of the protein than by a pure, aqueous medium.

Acknowledgment. This work was supported by a grant from the NASA Exobiology Program. C.C. was supported by a National Research Council Associateship. Computational resources from the National Aerodynamic Simulation facility (NAS) are gratefully acknowledged.

JA980010O