Early events in the folding of an amphipathic peptide

A Multi–Nanosecond Molecular Dynamics Study

Christophe Chipot[a], Bernard Maigret[a] and Andrew Pohorille[b,c]

[a] *Laboratoire de Chimie Théorique,*
Unité Mixte de Recherche CNRS No 7565,
Université Henri Poincaré–Nancy I, BP. 239,
54506 Vandœuvre-lès-Nancy Cedex — France

[b] *Exobiology Branch*
NASA — Ames Research Center
MS 239-4
Moffett Field, California 94035–1000

[c] *Department of Pharmaceutical Chemistry,*
University of California, San Francisco,
San Francisco, California 94143–0446

Correspondence to: Dr. Christophe Chipot, Laboratoire de Chimie Théorique, Unité Mixte de Recherche CNRS No 7565, Université Henri Poincaré–Nancy I, BP. 239, 54506 Vandœuvre-lès-Nancy Cedex — France. E-mail: chipot@lctn.u-nancy.fr

Abstract

Folding of the capped LQQLLQQLLQL peptide is investigated at the water–hexane interface by molecular dynamics simulations over 161.5 nanoseconds. Initially placed in the aqueous phase as a β–strand, the peptide rapidly adsorbs to the interface, where it adopts an amphipathic conformation. The marginal presence of non–amphipathic structures throughout the complete trajectory indicates that the corresponding conformations are strongly disfavored at the interface. It is further suggestive that folding in an interfacial environment proceeds through a pathway of successive amphipathic intermediates. The energetic and entropic penalties involved in the conformational changes along this pathway markedly increase the folding time–scales of LQQLLQQLLQL, explaining why the α–helix, the hypothesized lowest free energy structure for a sequence with a hydrophobic periodicity of 3.6, has not been reached yet. The formation of a type I β–turn at the end of the simulation confirms the importance of such motifs as initiation sites allowing the peptide to coalesce towards a secondary structure.

Introduction

Many proteins, involved in such essential cellular functions as energy transduction, signal transmission, transport of nutrients and waste products, cell fusion and narcosis, and catalysis of some metabolites, interact with cell membranes either by binding to their surfaces or by inserting themselves into the bilayer.[1,2] Integral membrane proteins are usually large and complex, and can contain both water-soluble and transmembrane regions. Whereas the former are usually built of diverse and flexible secondary structure elements, structural motifs characteristic of the latter are relative simple — *viz.* most often, either bundles of α-helices or β-barrels.[3] Typically, transmembrane elements of the secondary structure are quite stable, preserving their integrity in the membrane even if they are separated from the rest of the protein.[4-6] Moreover, the transmembrane segment alone, or with just a few additional residues, can retain the essential biological activity of the whole protein,[4,5,7,8] which is in sharp contrast with water-soluble proteins that only rarely preserve their structure or function upon truncation. This point is further reinforced by findings that some simple, synthetic peptides in the membrane can perform functions normally reserved to large proteins.[7,9,10] In view of the considerable difficulties in obtaining high resolution, three-dimensional structures of transmembrane proteins, this is very fortunate because it allows us to investigate simple peptide models with some faith in their biological relevance.

The structural stability of simple membrane peptides can also be used profitably to study protein folding, especially the early events leading to the formation of secondary structure elements. These events cannot be investigated conveniently using peptides dissolved in water, because only very few of them remain in a well-defined, ordered conformation. In contrast, membrane peptides, which are also usually disordered in water,[11-14] fold prior to, during or after insertion into the bilayer. Although the exact relation between folding and insertion has not been unequivocally established, most experimental[15] and theoretical[16-18] results support a hypothesis[15,19,20] that elements of the secondary structure are formed at the water-membrane

interface, and, only then, does the protein partition into the bilayer. Transmembrane proteins share this property of interfacial folding with peptides that reside at the membrane surface. These peptides are quite interesting in their own rights since they include many hormones, toxins, antibiotics and membrane fusion proteins.[21,22]

Similarities between transmembrane and surface peptides often extend much farther than folding behavior, and are manifested in the peptide sequences, many of which contain periodically spaced polar and non–polar amino acids. Possessing the appropriate hydrophobic periodicity, they can form α–helices, in which polar and non–polar residues are located at opposite faces. At the water–membrane interface, such an amphipathic helix can readily adopt an orientation in which the polar face is exposed to water, while the non–polar face is buried in the hydrophobic environment of the hydrocarbon tails in the core of the bilayer. This match between the polarities of the peptide and the environment should render the amphipathic α–helix particularly stable. Numerous experimental studies confirm that naturally occurring[23–26] and synthetic[13,27–31] peptides, as well as fragments of larger proteins,[32–35] which have sequences capable of forming amphipathic α–helices, indeed, do so. Typically, these peptides adopt orientations parallel to the interface.[26,36–39] At sufficiently high concentrations and/or in the presence of an electric field, some of them, however, rotate to the transmembrane orientation,[10,26,29,40] and may associate to form channel–like bundles with their polar faces lining the pore and their non–polar faces exposed to the hydrophobic environment in the membrane.

In this paper, we investigate the mechanism by which a simple peptide folds into an α–helical structure. In line with the above discussion, this is done at an interfacial environment, so that the results are relevant to both surface and transmembrane proteins. As a model, we have selected the Ac– and NHMe– terminally blocked undecamer built of L-leucine and L-glutamine residues. We have chosen the LQQLLQQLLQL sequence because it becomes amphipathic in the α–helical conformation. Although there are no experimental data for this particular peptide, a large body of evidence obtained for other amphipathic peptides, and discussed above, strongly

indicates that this molecule will form an α-helix at the interface. Particularly relevant to this work are studies on peptides similar to LQQLLQQLLQL, but built of L-leucine and L-serine, because the conformation of these peptides was shown to be α-helical at the water–membrane interface.[41] In this study, we have chosen L-glutamine rather than L-serine as the polar residue because the serine side chain cannot form easily hydrogen bonds with the peptide backbone which may importantly affect the folding of many amphipathic peptides. Interestingly enough, as has been shown by O'Neil and DeGrado,[42] L-glutamine and L-serine are equivalently good helix formers.

Our approach is based on atomic–level molecular dynamics simulations with the explicit representation of the solvent. The main difficulties of this approach are the limited time–scales accessible to such calculations. For this reason, computer simulations of the mechanism of peptide folding were limited, until recently, to approximate approaches that relied on an implicit representation of the solvent,[43–45] and a simplified description of the amino acid residues.[46–48] Molecular dynamics simulations with full atomic representation of both the protein and the solvent were previously employed only to study unfolding,[49–51] often at unrealistically high temperatures chosen to accelerate this process. Only lately, advances in computer technologies have made it possible to investigate folding directly using the same methodology.[14,52–54] Most notably, pathways to folding of the 36 residue–long villin headpiece subdomain in water into a state that resembles the native conformation were observed in a 1 microsecond simulation.[52] This study required a very large commitment of computational resources of a massively parallel supercomputer, clearly demonstrating that even relatively simple folding problems remain a challenge.

For systems at the water–membrane interface, additional factors complicate the simulations. One of these is the slow relaxation of lipid molecules in response to conformational changes of the peptide. Furthermore, the high charge density in the phospholipid head group region makes it necessary to include explicitly long-range electrostatic effects. Finally, standard molecular

dynamics methods may not be appropriate for systems in which the folding of a protein is accompanied by changes in the dimensions of the simulation box. Instead, an "extended system" treatment may be required.[55,56] The dynamics of the protein in the extended system, however, is no longer Newtonian, and it is not known to what extent it captures the real dynamics of folding.

Many of the difficulties outlined above can be avoided by considering a water–alkane, instead of a water–membrane, system. Such a membrane–mimetic system captures the most important characteristic of water–membrane interfaces — the coexistence of a polar, aqueous phase and a non–polar medium. In fact, water–alkane or even water–air interfaces have been used occasionally in both experimental[34,41,57-60] and theoretical[14,61-64] studies of membrane–active peptides. In particular, it has been shown that amphipathic peptides fold into the same secondary structure at all three interfaces.[41,57,58] These considerations motivate our choice of water–hexane rather than water–membrane as the interfacial system. It should be, however, kept in mind that such an approximation has important limitations due to the absence of the many complexities of the water–membrane interfacial region.[33,65,66] In particular, electrostatic interactions between the peptide and the lipid head groups are neglected, even though they have been shown to affect folding of membrane surface proteins, especially if the protein contains charged residues.[67]

The present work can be considered as an extension of our molecular dynamics study on the folding of the undecamer of poly–L–leucine at the water–hexane interface.[14] In that study, the peptide fully folded into an α–helix in approximately 36 ns. In the simulation presented here, complete folding did not occur even though the molecular dynamics trajectory exceeded 160 ns. The simulation, nevertheless, sheds an interesting light on the early events of interfacial folding. Its results and their discussion are presented after the section covering the methodological aspects of the study. The paper closes with a summary and conclusions relevant to systems of biological interest.

Model and methods

The molecular dynamics simulations were carried out in the microcanonical (N, V, $\mathscr{E}$) ensemble, using the program COSMOS.[68] The interfacial environment of the Ac– and –NHMe terminally blocked LQQLLQQLLQL undecapeptide consisted of a water lamella of 1367 molecules in contact with a hexane lamella of 284 molecules. Each lamella was in equilibrium with its vapor phase. The dimensions of the simulation cell were 45.0 × 45.0 Å^2 in the x,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 spatial directions. The Newton equations of motion were integrated using the Verlet algorithm,[69] with a time step of 2 fs. The length of the molecular dynamics trajectory generated to monitor the interfacial folding of the undecapeptide amounted to 161.5 ns. In order to maintain the temperature of the system at 300 K, the velocities of all the atoms were occasionally rescaled. Such rescaling, necessary to correct for inaccurate integration of the equations of motion, was performed only if during a 10 ps time period the average temperature departed from the required value by more than ± 3 K. Over 100 ps of MD trajectory, the deviation in the total energy was, on average, 0.44 %.

The water molecules in the system were represented by the TIP4P model.[70] The hexane molecules were described by the OPLS potential energy functions,[71] in which methyl and methylene groups are represented by single, united atoms of the appropriate radius. It has been shown that this simple description of aliphatic chains yields good estimates of interfacial thermodynamic quantities such as the water–alkane surface tension,[72] or the water–hexane partition coefficients of small, organic molecules.[73,74] Although a united atom model generally appears to be sufficient for reproducing thermodynamic equilibrium properties, it may lead to overestimated diffusion rates in the non–polar phase, and, thus, can somewhat influence folding times. All intramolecular parameters of the potential energy function describing the undecapeptide were taken from the AMBER all–atom force field of Cornell et al.[75] Parameters representing the water–peptide and the hexane–peptide interactions were computed from the standard OPLS

combination rules.[71]

Pairwise Coulomb and van der Waals intermolecular interactions involving water molecules and/or small, electrically neutral groups[76] in the peptide (*e.g.* $-CH_3$, $-C=O$, or $-NH$ moieties) or in hexane molecules were smoothly truncated between 7.5 and 8.0 Å, by means of a cubic spline switching function applied to both the energy and the forces.[77,78] The truncation was applied to the group centers. Long–range, electrostatic interactions were not taken into account in this simulation. This is an acceptable approximation, because the water–hexane interface is lacking charged or zwitterionic lipid head groups present in the water–membrane system, and the terminally blocked LQQLLQQLLQL is uncharged and not strongly polar.[79]

High frequency degrees of freedom were eliminated by constraining the bond lengths and the valence angles of both water and hexane molecules, and the bond lengths between heavy atoms and hydrogen atoms in the undecapeptide to their equilibrium values. This was done using the SHAKE algorithm.[80]

The simulation was performed on a single R10000, 195 MHz processor of a 10-processor SILICON GRAPHICS *Power Challenge*. The average CPU time per one molecular dynamics step was 0.51 second, yielding a total CPU time of 476 days for the complete trajectory.

Results and discussion

Initially, the peptide was placed in the aqueous medium, near the water–hexane dividing surface. Its starting conformation was a β–strand, which, for the sequence chosen, is not amphipathic. The undecapeptide, however, rapidly adsorbed to the interface and its backbone underwent conformational transitions that yielded an amphipathic structure, wherein the leucine side chains were buried in hexane and the glutamine side chains were immersed in water. As the density profiles in Figure 1 indicate, the peptide, once adsorbed at the interface, remained in this environment for the rest of the simulation. It exhibited a slight preference for the

hexane side, perhaps due to the small imbalance in the number of hydrophobic and hydrophilic residues. The probability distribution, $\mathscr{P}(z)$, of finding the center of mass of the peptide at a given distance, z, from the interface was converted to a free energy profile, $\Delta A(z)$, along the z-direction, using the equation:

$$\Delta A(z) = -k_B T \log \mathscr{P}(z) \tag{1}$$

where k_B is the Boltzmann constant and T is the temperature. The resulting curve, shown in Figure 2, clearly indicates that the undecapeptide is interfacially active — i.e. tends to accumulate at the interface — and suggests that there is a significant free energy penalty for its desorption into one of the bulk phases. As can be seen from Figures 1 and 2, however, the center of mass of the peptide is not rigidly pinned to the interface, but instead undergoes considerable translational motions along z. A similar result was obtained for the undecamer of poly-L-leucine at the same interface.[14]

To analyze the structural relationships among the conformations adopted by the peptide during the simulation, the distance root mean square deviations (RMSDs) for all pairs of conformations were calculated using backbone atoms only. The results are displayed in Figure 3 as a two-dimensional graph. In this graph, every pair for which the RMSD was less than 2 Å is plotted as a dot. The clustering analysis indicates the presence of six distinct structural families, denoted (a)-(f), which exist in the following time, t, intervals:

$$
\begin{aligned}
\text{(a)} & \begin{cases} 0.0 & \leq t \leq & 6.2 \text{ ns} \\ 10.1 & \leq t \leq & 90.0 \text{ ns} \\ 103.1 & \leq t \leq & 108.0 \text{ ns} \end{cases} \\
\text{(b)} \quad & 6.3 \leq t \leq 10.0 \text{ ns} \\
\text{(c)} \quad & 90.1 \leq t \leq 103.0 \text{ ns} \\
\text{(d)} \quad & 108.1 \leq t \leq 121.0 \text{ ns} \\
\text{(e)} \quad & 121.1 \leq t \leq 149.0 \text{ ns} \\
\text{(f)} \quad & 149.1 \leq t \leq 161.5 \text{ ns}
\end{aligned}
$$

Whereas the first three structural families are somewhat diffuse and show some similarities to

each other, each of the last three families appears to show little resemblance to any of the remaining five families. Typical conformations from the six families identified in the clustering analysis are shown in Figure 4.

The two-dimensional graph, shown in Figure 3, indicates that the peptide is not conformationally rigid, but slowly evolves in conformational space. From this graph, however, one cannot learn whether the folding pathway of the peptide actually leads to a helical structure. This information can be obtained, nonetheless, by correlating the results of the graph with the Ramachandran maps for the (ϕ, ψ) dihedral angles of the 11 residues forming the undecapeptide. This is done in Figure 5. In these plots, the size of the points characterizing the conformation of a given residue corresponds to a particular structural family emerging from the clustering analysis — *i.e.* the smallest, darkest dots denote conformations of family (a), whereas the largest, lightest dots represent conformations of family (f). As can be seen, the degree of flexibility exhibited by the 11 peptide bonds along the backbone differs. Four residues, *viz.* Leu^1, Gln^2, Leu^4 and Leu^9, remain confined to a single quadrant of the (ϕ, ψ) map throughout the simulation. Three of them, Leu^1, Gln^2 and Leu^9, occupy the "β" region, oscillating between a β-strand and an extended conformation. For Leu^4, only the right-handed, "α_R" quadrant is sampled, indicating that, early in the simulation, this residue moved away from its starting conformation to ensure the amphipathy of the undecapeptide. Residues Gln^3, Leu^5, Leu^8, Gln^{10} and Leu^{11} explore both the "β" and the "α_R" quadrants, but eventually return to the former. Only residues Glu^6 and Glu^7 sample the two regions before adopting an α_R conformation at the end of the simulation.

The analysis of the Ramachandran maps indicates that, although the peptide adopts different structures from several distinct families, none of them resembles closely an ordered conformation, such as an α-helix. It, therefore, cannot be simply characterized by just a few parameters. Instead, these structures can be well described only by a set of global and local indices that complement the maps in Figure 5.

One such global index that provides a comprehensive, characterization of changes in the shape of short peptides during folding can be obtained from their radii of gyration, $\{R_0^k\}$, defined as:

$$R_0^k = \sqrt{\frac{I^k}{M}} \qquad k = 1, 2, 3 \qquad (2)$$

where $\{I^k\}$ denote the eigenvalues of the instantaneous inertia tensor of the peptide and M is its total mass. From the set $\{R_0^k\}$, a parameter, A_3, referred to as asymmetry of the peptide, can be derived:[81]

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

The asymmetry, A_3, can take values between 0 and 1, with 0 corresponding to a perfectly spherical shape, and 1 to a linear one. In Figure 6, the time evolution of A_3 is displayed and compared to the values representative of the ideal β–strand, 3_{10}- and α–helices.[82] For the first 108 ns, when the peptide evolves through structural families (a)–(c), its shape remains close to that of the β–strand, which was the starting point of the simulation. This global characteristic agrees with the impression from Figure 4. In the subsequent 40 ns, as the peptide adopts conformations from structural families (d) and (e), the asymmetry decreases and fluctuates around the value representative of an α–helix. Since many different conformations may correspond to the same value of A_3, this similarity does not imply that the peptide becomes helical. It, however, suggests that families (d) and (e) may exhibit some structural resemblance to an α–helix. This expectation is, indeed, confirmed by examining the formation of intra–backbone hydrogen bonds between the carbonyl group of residue $i - 4$ and the amino group of residue i, which constitute the scaffolding of the α–helix. As may be observed from the time–evolution of all the nine $C{=}O_{i-4}\cdots H{-}N_i$ distances in the peptide, shown in Figure 7, three of them, $viz.$ distances 5, 6 and 7, decrease to approximately 2 Å between 108 and 120 ns. This clearly

points to the formation of three consecutive hydrogen bonds along the backbone near the center of the peptide. These bonds, however, are not rigid but exist in an equilibrium with the $C=O_{i-3}\cdots H-N_i$ hydrogen bonds characteristic of a 3_{10}-helix.

The impression that the peptide begins to adopt a helical structure is further reinforced by examining Figure 8, which shows the time evolution of the distance RMSD between the undecapeptide and ideal 3_{10}- and α-helices.[82] For the first 80 ns, average values of both RMSDs remain approximately constant. Upon the transition to family (d), the RMSD with respect to the α-helix, however, decreases substantially, by nearly a factor of two. For the next 40 ns, again, the RMSD only fluctuates around an average value, which is close to the RMSD with respect to the 3_{10}-helix. This similarity reflects an apparent dynamic inter-conversion between the two types of helices in the folded fragment of the peptide.

Although the nature of structures in families (d) and (e) appears to indicate that the peptide progresses towards the α-helical state, further time-evolution of the system demonstrates that this is not the case. After 149 ns, all $C=O_{i-4}\cdots H-N_i$ distances become longer than 6 Å, as they were at the beginning of the simulation. Simultaneously, the asymmetry decreases to nearly zero, indicating that the peptide adopts a more spherical shape, and the distance RMSD with respect to the α-helix increases to approximately 6 Å. All these indices point to a substantial conformational change that accompanies the transition from the structural family (e) to family (f). The snapshot of a representative conformation from the latter family, shown in Figure 4, reveals the nature of this change — the peptide adopts a hairpin conformation with a structure similar to a type I β-turn formed in the middle of the sequence. Since this type of a reverse turn is just a deformed 3_{10}-helix, it is expected that a hydrogen bond be present in the middle of the backbone. This is, indeed, the case: The fifth $C=O_{i-3}\cdots H-N_i$ distance in the undecapeptide becomes equal to approximately 2 Å and fluctuates around this value until the end of the simulation. This clearly points to the presence of the required hydrogen bond, supporting the identification of the structure as a β-turn.

The observed transition raises a question: Why did the peptide unfold instead of continuing a cooperative folding to an α-helix? It should be noted that three consecutive backbone hydrogen bonds, present in families (d) and (e) were located in the middle of the peptide rather than at one end, as would be expected in standard theories of helix-coil transitions in bulk media.[83,84] One may speculate that this is precisely the reason for unfolding — the formation of an α-helix started from a wrong place and the peptide has to refold completely before it starts folding again, presumably from the end. Based on experience from the simulation of the folding of the terminally blocked undecamer of poly-L-leucine at the water-hexane interface,[14] this is not a likely explanation. Poly-L-leucine folded completely to a helical structure, but along a pathway that did not start from the end of the peptide and was not sequential. In fact, there is another interesting similarity between the two pathways — approximately half-way through the folding process, the already formed hydrogen bonds in poly-L-leucine were intermittently broken to create a turn in the peptide. This disordering was also accompanied by an increase in the distance RMSD with respect to the helical structures, and a large decrease in A_3. The turn, however, was only a transient state, and, subsequently, the broken hydrogen bonds reformed, permitting further folding. As commented on recently by Dobson et al.,[85] β-turns or nascent helices can be viewed as initiation sites that increase the probability to coalesce towards motifs involving longer range contacts. By analogy, it can be argued that the β-turn in LQQLLQQLLQL is also an intermediate, and, in a longer simulation, the peptide would return to its partially folded state. This would be consistent with the concept of the nascent helix, an ensemble of transient, partially ordered secondary structures, which rapidly inter-convert by way of unfolded states.[86] The study on poly-L-leucine indicates that this concept applies to interfacial folding.[14] The presence of a β-turn among conformations of the nascent helix is not surprising. Several amphipathic peptides were shown to adopt this structure at the water-membrane interface.[87-90] Furthermore, in water-soluble proteins, β-turns exist predominantly at protein surfaces, in contact with the solvent.[91] It has also been argued that β-turns are common folding intermediates.[92] It is, thus, likely that their stability at an interface may not

be very different from the stability of short, partially folded helices.

Although the folding of the undecamer of poly–L–leucine and LQQLLQQLLQL exhibit some similarities, there are also important differences, the chief of which is the time–scales of these two processes. Poly–L–leucine fully folded in 36 ns, whereas LQQLLQQLLQL partially folded after over 100 ns, only to unfold towards the end of the 161 ns simulation. Although single trajectories are insufficient to draw conclusions about time–scales of folding events, the difference in these time–scales for the two peptides is sufficiently large to infer that LQQLLQQLLQL folds more slowly than poly–L–leucine. To understand why this might be so, even though simulation conditions and potential energy functions were the same in both cases, it is necessary to consider how various characteristics that differentiate the two peptides may influence folding.

Before we proceed to the discussion of this point, we first consider the hypothesis that an α–helix is the most stable structure for LQQLLQQLLQL. Although there is no experimental data for this peptide, results for other amphipathic peptides, presented in the introduction, support this assumption. A thorough test in computer simulations, at least to within the accuracy of the potential energy functions, would require calculating the free energy of the α–helix relative to other conformational states of the peptide. This daunting task has not been carried out in this study. Instead, a separate, 11.3 nanosecond–long molecular dynamics trajectory was obtained starting with an equilibrated α–helical conformation of the undecapeptide placed at the water–hexane interface. For the entire length of the simulation, the peptide remained in an amphipathic orientation, such that the leucine and the glutamine side chains were immersed in hexane and water, respectively. The distance RMSD between the simulated peptide and the ideal helix remained quite small throughout the trajectory, and was equal, on average, to 0.53 Å. Similarly, A_3 was equal to 0.172, close to the value of the ideal α–helix, with the standard deviation of 0.003. This deviation is much smaller than the standard deviation for the peptide in the 161.5 ns simulations, equal to 0.045, This, in turn, indicates that the α–helix is markedly more rigid than structures in the families (a–f), and exhibits no tendency

to melt, which would be manifested by an increased flexibility. Although the length of this simulation is not sufficient to draw a firm conclusion that an α-helix is the preferred structure at the interface, the remarkable stability of this conformation during the simulation certainly supports such an assumption.

Another view on the stability of the α-helix can be obtained by calculating the free energy of unfolding a single residue at both the N- and the C-terminus ends of the undecapeptide — i.e. the free energy for going from a right-handed α-helical to a β-strand conformation.[93] The starting point for this series of calculations was an α-helical conformation at the water-hexane interface obtained at the end of the 11.3 ns simulation described above. The free energy of unfolding was evaluated as a function of the ψ dihedral angle, using the "umbrella sampling" method.[94] In the case of the N-terminus, five overlapping "windows"[95,94] were needed to explore the range of values accessible to ψ, i.e., $-60.0 \leq \psi \leq 160.0$. At the C-terminus, four windows were necessary to examine the same range of ψ. The lengths of sampling trajectories in each window varied between 1.2 and 3.0 ns, yielding a total simulation time of 11.5 and 7.0 ns to estimate the free energy of unfolding at the N- and the C-terminus, respectively. To restrain the values of ψ in a given window, a harmonic potential was applied. Furthermore, a biasing potential was added to yield a more uniform probability distribution of the dihedral angle within each window, thereby improving the accuracy of the calculations. The complete free energy profiles at both ends were obtained by matching in a self-consistent manner the individual curves in the overlapping regions, using the weighted histogram analysis method (WHAM).[96] During this set of simulations, the remainder of the undecapeptide was restrained in an α-helical conformation by means of soft harmonic potentials.

As can be observed from Figure 9, the α-helix is considerably more stable than both conformations unfolded at the ends. The two termini, however, are not equivalent. The free energy difference between the α-helical and the β-strand conformation of the peptide bond is markedly smaller at the C-terminus than at the N-terminus, viz. 4.75 and 8.66 kcal/mol,

respectively. This suggests that the peptide structure is more flexible at the N–terminus than at the C–terminus. This agrees with the Ramachandran maps for residues Leu[1] and Leu[11], shown in Figure 5 — Leu[1] stays in the "β" quadrant over the whole 161.5 ns of the simulation, while Leu[11] explores both the "α_R" and the "β" quadrants. Moreover, the free energy barrier for rotation from α_R to β is approximately 10 kcal/mol. If barriers between other conformational states are similar, it may explain why the undecapeptide evolves very slowly in its conformational space at 300 K. The results from the calculations presented here should be taken cautiously, however. Simulations of unfolding of poly–L–leucine from the N–terminus in water,[93] using the Cornell *et al.* potential energy functions[75] also yielded a high free energy difference between the folded and unfolded states, together with a significant free energy barrier. Yet, once the potentials were modified to reproduce relative energies of different conformations of short peptides in vacuum, obtained from *ab initio* quantum mechanical calculations,[97] both the free energy differences and the barriers were markedly reduced.

Returning to the issue of slow conformational evolution, we note that one structural feature distinguishing LQQLLQQLLQL from poly–L–leucine is the ability of the glutamine side chains to form intramolecular hydrogen bonds with the polar moieties of the backbone. Once such bonds are formed, the hydrogen–bonded centers in the backbone are no longer available for the through–backbone interactions necessary to stabilize helical conformations. The time–evolution of the number of backbone–side chain hydrogen bonds is displayed in Figure 10. As can be seen from this figure, such bonds are almost always present and, occasionally, as many as four of them exist at the same time. The most common are those formed between the $-NH_2$ group of the side chain and the carbonyl group of the backbone in the adjacent residue. In contrast, bonds to the carbonyl group of the same residue are scarce. Hydrogen bonds between the $-C=O$ moiety of the glutamine side chain and the amino groups of the backbone are equally marginal. Backbone–side chain hydrogen bonds frequently exist at the N–terminus of the peptide, but only rarely at the C–terminus. Specifically, 39% of them form between $-NH_2$ moiety of Gln[1]

and the carbonyl of the –Ac capping group, and an additional 24% is formed between the –NH$_2$ group of Gln^3 and the carbonyl of Leu^2. A similar tendency towards the formation of backbone–side chain hydrogen bonds at the capping positions was also noted for water–soluble proteins.[98] In the context of the present simulation, these hydrogen bonds may explain why the formation of a partial helix started near the middle of the peptide rather than at the C–terminus.

A similar analysis carried out on the 11.3 ns molecular dynamics trajectory for the α–helical undecapeptide reveals that the backbone–side chain hydrogen bond population is marginal and involves only Gln^2 and Gln^3. In contrast with the results obtained from the 161.5 ns trajectory, only hydrogen bonds between the –NH$_2$ moiety in the side chain and the carbonyl group of the same residue are observed. This, again, supports the assumption than an amphipathic α–helix of LQQLLQQLLQL is a stable structure, at least over short time–scales.

Another factor that may influence the kinetics of α–helix formation is connected with the nature of folding intermediates formed by LQQLLQQLLQL. The simulation of poly–L leucine at the water–hexane interface revealed that its folding proceeded through a short–lived intermediate, the 3_{10}–helix, directly preceding the formation of the α–helix.[14] This is in agreement with experimental studies on short poly–L–alanine peptides in aqueous solution carried out by Millhauser et al.,[99,100] which led to the conclusion[101] that a 3_{10}–helix is an important intermediate on the pathway connecting a random coil to an α–helix. A similar conclusion emerged from other experimental[92] and theoretical studies.[102–104] The exact role of 3_{10}–helices in facilitating folding is not clear, but it may be important that the formation of a through–backbone hydrogen bond in this structure requires that two consecutive peptide bonds adopt appropriate (ϕ, ψ) angles, while three peptide bonds must have the correct conformation to form a hydrogen bond in an α–helix. Once a 3_{10}–helix is formed, conversion to an α–helix should be highly facilitated, both kinetically and thermodynamically. In fact, several peptides have been shown to exist in an equilibrium between the two types of helices both in water[105] and in non–polar media.[106–110] For LQQLLQQLLQL and other peptides capable of forming amphipathic α–helices,

3_{10}-helices are, however, not amphipathic, and, therefore, may not be stable intermediates at interfaces. This inability to reach the 3_{10}-helix on the folding pathway may, in turn, impede folding.

The suggestion that a non-amphipathic conformation of the 3_{10}-helix may not be a folding intermediate raises the question whether the structures encountered during the simulation are amphipathic. The density profiles in Figure 1 indicate that this is, indeed, the case. To explore this point further, the degree of amphipathy of LQQLLQQLLQL was estimated from its mean structural hydrophobic moment,[111] $\langle \mu_H \rangle$, computed over the complete trajectory:

$$\langle \mu_H \rangle = \frac{1}{N} \sum_{i=1}^{N} h_i\, s_i \tag{4}$$

where N denotes the number of amino acid residues in the peptide, h_i is the hydrophobicity of residue i, and s_i is the unit vector pointing from the α-carbon atom of residue i to the center of mass of its side chain. Hydrophobicities, h_i, were taken from the set published by Wolfenden, et al.,[112] but, as observed by Eisenberg et al., $\langle \mu_H \rangle$ does not depend significantly on the choice of the hydrophobicity scale. In Figure 11, the time-evolution of $\langle \mu_H \rangle$ is shown and compared to the values characteristic of the ideal β-strand, 3_{10}- and α-helices.[82] The most striking feature of this graph is that the average of $\langle \mu_H \rangle$ over the full MD trajectory is equal to 0.32, which is very close to the value of 0.34, expected for an α-helix, but quite different from 0.20 and 0.16, typical to a β-strand or a 3_{10}-helix, respectively. It, thus, appears that, although the undecapeptide is not in an α-helical conformation, its amphipathy at the water-hexane interface is nearly optimal.

The results, discussed above, lead to a general hypothesis about interfacial folding of LQQL-LQQLLQL and other peptides with similar periodicity of hydrophobic and hydrophilic amino acids. All of them tend to adopt amphipathic structures. The α-helix is probably the most stable of them, but there are many others. Thus, driven by the tendency to maintain orienta-

tions in which hydrophobic and hydrophilic residues are exposed, respectively, to the aqueous and the non-polar phases, peptides evolve in their conformational space along amphipathic pathways. Transitions between consecutive amphipathic structures may, however, require surmounting substantial free energy barriers associated to the dehydration of polar moieties. In addition, folding may be further frustrated by bottlenecks created by the presence of non-amphipathic conformations on folding pathways. This would yield a frustrated free energy surface with multiple, deep, local minima that would impede folding by acting as "kinetic traps." A qualitatively similar mechanism has been proposed to explain folding of large, water-soluble proteins.[113]

Summary and Conclusions

The 161.5 ns molecular dynamics simulation of LQQLLQQLLQL in the water–hexane system reveal. the importance of amphipathic structures in defining conformational transitions of t.is peptide at interfaces between water and non-polar media. Many such structures are available in conformational space and some of them were explored during the simulation.

The almost complete absence of non-amphipathic conformations suggests that they are strongly disfavored at the interface. On this basis, it is hypothesized that the undecapeptide and other, similar, short peptides fold in the interfacial environment through a series of amphipathic intermediates. In the present simulation, however, the α–helix, which presumably is the lowest free energy structure, was never reached. It is further hypothesized that energetic and entropic penalties associated with folding along amphipathic pathways markedly impede evolution of the peptide in its conformational space.

Several other factors may be also responsible for slow folding of LQQLLQQLLQL, compared to poly-L-leucine, which adopted an α-helical conformation in less than 40 ns.[14] One of them is the inability of LQQLLQQLLQL to pass through a 3_{10}-helical folding intermediate, which is not

amphipathic for this peptide. Another is the formation of hydrogen bonds between the side chains of the L-glutamine residues and the carbonyl groups in the backbone. The significance of these bonds might be established more clearly by comparing, for example, folding pathways for LQQLLQQLLQL and a similar undecapeptide with L-serine substituted for L-glutamine.

The peptide initially remained in conformations resembling the starting configuration — i.e. a β-strand. Subsequently, it partially folded, forming three consecutive backbone hydrogen bonds characterisitic of an α-helix, only to unfold into a type I β-turn towards the end of the simulations. By analogy to the simulated folding of poly-L-leucine at the water-hexane interface,[14] it may be proposed that these conformations are not just typical structures in the ensemble of unfolded states, but represent the nascent helix, which will eventually evolve to the fully folded conformation.

Interfacial folding of amphipathic peptides is of interest not only in the context of water-membrane systems, but may also be relevant to folding of water-soluble proteins. In those instances, amphipathic helices do not form fully exposed to water, but rather at the "interface" between the solvent and the rest of the protein, accompanied by its hydrophobic collapse. This point was illuminated in the recent molecular dynamics simulation of protein folding in an aqueous solution[52] and is further underscored by findings that some peptides which are disordered in water in their monomeric form adopt α-helical conformations upon aggregation.[114,115]

To evaluate results of this study one should be cognizant of its limitations. One of them is associated with the timescale of the simulation. Even though the molecular dynamics trajectory discussed in this study is among the longest obtained for biologically relevant molecules in condensed phases, it is clearly not sufficiently long. As has been shown in the example of a small protein in aqueous solution,[52] extending the trajectory by an order of magnitude is feasible. Yet, even if the peptide folded completely, as was the case for poly-L-leucine at the water-hexane interface, it still would not be fully satisfactory, since drawing far-reaching

conclusions from a single trajectory is not justified. Instead, a statistical sample of folding pathways should be obtained, but this is beyond reach of current computer capabilities. Along these lines, the present simulation should not be treated as an isolated study, but rather as a complement to a growing body of experimental and theoretical knowledge about interfacial folding.

Another limitation of the present study is related to the quality of potential energy functions. These functions have been calibrated to reproduce either free energy surfaces of very small protein fragments or equilibrium structures of water-soluble proteins. Relatively little effort, however, has been devoted to parametrizing potential energy functions such that they would correctly predict free energy barriers during conformational transitions in peptides and proteins.

Finally, the water-hexane interface is only a very simplified model of a water-membrane system. Several effects that may influence folding, such as specific and non-specific interactions between the peptide and the head groups, reorganization of the bilayer and matching between the width of the peptide and the head group region, cannot be captured using this model.

Despite all these difficulties, atomic-level computer simulations remain a promising approach to protein folding in solution. Considering the great technical difficulties in experimental investigations of early folding events at water-membrane interfaces, such as the formation of secondary structures and insertion into the bilayer, these computational approaches become especially valuable. In the near future, they are anticipated to provide many valuable insights into the mechanism of interfacial folding.

Acknowledgment

Figure captions

Figure 1. Density profiles of water (solid line), hexane (long–dashed line), the center of mass of the undecapeptide (short–dashed line), leucine side chains (dotted–dashed line) and glutamine side chains (dotted line) at the water–hexane interface averaged over the complete molecular dynamics trajectory.

Figure 2. Free energy profile for finding the center of mass of LQQLLQQLLQL along the z–direction perpendicular to the interface. This curve was constructed using the last 71.5 ns of the molecular dynamics trajectory.

Figure 3. Two–dimensional distance root mean square deviation (RMSD) map for all pairs of structures of the LQQLLQQLLQL undecapeptide encountered during the 161.5 ns trajectory. Only those pairs for which the distance RMSD was less than 2.0 Å are plotted as points on the map. Distance RMSDs were computed for backbone atoms only.

Figure 4. Snapshots of typical conformations of the undecapeptide in each of the families (a)–(f) identified on the basis of the clustering analysis in Figure 3. Intramolecular hydrogen bonds are represented by dotted lines.

Figure 5. (ϕ, ψ) Ramachandran maps for the eleven residues forming the terminally blocked undecapeptide. Conformations within the same structural family are represented by dots of the same size and hue. The size of dots increases along the trajectory, $i.e.$, the smallest, darkest dots represent family (a), whereas the largest, lightest dots represent family (f).

Figure 6. Time–evolution of the asymmetry, A_3, of the undecapeptide (solid line). Characteristic values of A_3 for the α–helix (short–dashed line), the β–strand (long–dashed line) and the 3_{10}–helix (dotted line) are also shown.

Figure 7. Time–evolution of the nine $C=O_{i-4}\cdots H-N_i$ intramolecular backbone hydrogen bonds

of the undecapeptide.

Figure 8. Distance root mean square deviation (RMSD) of the LQQLLQQLLQL undecapeptide simulated over 161.5 ns at the water–hexane interface, with respect to an ideal α–helix, *i.e.* $\phi = -57°, \psi = -47°$ (solid line) and 3_{10}–helix, *i.e.* $\phi = -49°, \psi = -26°$ (dashed line).[82]

Figure 9. Free energy profile for unfolding the first, N–terminal (solid line) and the last, C–terminal (dashed line) residues in the α–helical structure of the terminally blocked LQQL-LQQLLQL undecapeptide at the water–hexane interface.

Figure 10. Time–evolution of the number of intramolecular backbone–side chain hydrogen bonds in the undecapeptide at the water–hexane interface.

Figure 11. Time–evolution of the mean structural hydrophobic moment $\langle \mu_H \rangle$ (solid line) and the average of $\langle \mu_H \rangle$ (long–dashed line) of the undecapeptide. The values characteristic of the α–helix (short–dashed line), the β–strand (dotted line) and the 3_{10}–helix (dotted–dashed line) are also shown.

References

[1] Gennis, R. B., *Biomembranes: Molecular Structure and Function*, Springer: New York, 1989.

[2] Petty, H. R., *Molecular Biology of Membrane Structure and Function*, Plenum Press: New York, 1993.

[3] Cowan, S.; Rosenbusch, J., Folding pattern diversity of integral membrane proteins, *Science* **1994**, *264*, 914–916.

[4] Oblatt-Montal, M.; Buhler, L.K.; Iwamoto, T.; Tomich, J.M.; Montal, M., Synthetic peptides and four-helix bundle proteins as model systems for the pore–forming structure of channel proteins. i. Transmembrane segment M2 of the nicotinic cholinergic receptor channel is a key pore–lining structure, *J. Biol. Chem.* **1993**, *268*, 14601–14607.

[5] Reddy, G.L.; Iwamoto, T.; Tomich, J.M.; Montal, M., Synthetic peptides and four–helix bundle proteins as model systems for the pore-forming structure of channel proteins. II. Transmembrane segment M2 of the brain glycine receptor is a plausible candidate for the pore–lining structure, *J. Biol. Chem.* **1993**, *268*, 14608–14615.

[6] Corbin, J.; Méthot, N.; Wang, H.H.; Baenziger, J.E.; Blanton, M.P., Secondary structure analysis of individual transmembrane segments of the nicotinic acetylcholine receptor by circular dichroism and Fourier transform infrared spectroscopy, *J. Biol. Chem.* **1997**, *273*, 771–777.

[7] Montal, M., Molecular mimicry in channel–protein structure, *Curr. Opin. Struct. Biol.* **1995**, *5*, 501–506.

[8] Duff, K. C.; Ashley, R. H., The transmembrane domain of influenza A M2 protein forms amantidine sensitive proton channels in planar lipid bilayers, *Virology* **1992**, *190*, 485–489.

[9] Sansom, M. S. P., The biophysics of peptide models of ion channels, *Prog. Biophys. Mol. Biol.* **1991**, *55*, 139–235.

[10] Lear, J. D.; Schneider, J. P.; Kienker, P. K.; DeGrado, W. F., Electrostatic effect on ion selectivity and rectification in designed ion channel peptides, *J. Am. Chem. Soc.* **1997**, *119*, 3212–3217.

[11] Vogel, H., Incorporation of melittin into phosphatidylcholine bilayers, *FEBS Lett.* **1981**, *134*, 37–42.

[12] White, S. H.; Wimley, W. C.; Ladokhin, A. S.; Hristova, K., Protein folding in membranes: Determining energetics of peptide–bilayer interactions, *Methods Enzymol.* **1998**, *295*, 62–87.

[13] Tamm, L. K.; Tomich, J. M.; Saier, M. H., Membrane incorporation and induction of secondary structure of synthetic peptides corresponding to the N-terminal signal sequences of the glucitol and mannitol permeases of *Escherichia coli*, *J. Biol. Chem.* **1989**, *264*, 2587–2592.

[14] Chipot, C.; Pohorille, A., Folding and translocation of the undecamer of poly–L–leucine across the water–hexane interface. A multi–nanosecond molecular dynamics study, *J. Am. Chem. Soc.* **1998**, *120*, 11912–11924.

[15] Jacobs, R. E.; White, S. H., The nature of the hydrophobic binding of small peptides at the bilayer interface: Implications for the insertion of transbilayer helices, *Biochemistry* **1989**, *28*, 3421–3437.

[16] Milik, M.; Skolnick, J., Spontaneous insertion of polypeptide chains into membranes: A Monte Carlo model, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9391–9395.

[17] Milik, M.; Skolnick, J., Insertion of peptide chains into lipid membranes: An off–lattice Monte Carlo dynamics model, *Proteins* **1993**, *15*, 15–25.

[18] Ben-Tal, N.; Ben-Shaul, A.; Nicholls, A.; Honig, B., Free-energy determinants of α-helix insertion into lipid bilayers, *Biophys. J.* **1996**, *70*, 1803–1812.

[19] Engelman, D.; Steitz, T. A., The spontaneous insertion of proteins into and across membranes: The helical hairpin hypothesis, *Cell* **1981**, *23*, 411–422.

[20] Popot, J. L.; Gerchman, S. E.; Engelman, D. M., Refolding of bacteriorhodopsin in lipid bilayers: A thermodynamically controlled two-stage process, *J. Mol. Biol.* **1987**, *198*, 655–676.

[21] Kaiser, E. T.; Kedzy, F. J., Peptides with affinity for membranes, *Ann. Rev. of Biophys. Biophys. Chem.* **1987**, *16*, 561–581.

[22] Watts, A., Solid-state NMR apporaches for studying the interaction of peptides and proteins with membranes, *Biochim. Biophys. Acta* **1998**, *1376*, 297–318.

[23] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A., Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation, *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.

[24] Pérez-Payá, E.; Dufourcq, J.; Braco, L.; Abad, C., Structural characterisation of the natural membrane-bound state of melittin: A fluorescence study of dansylated analogue, *Biochim. Biophys. Acta* **1997**, *1329*, 223–236.

[25] Bernèche, S.; Nina, M.; Roux, B., Molecular dynamics simulation of melittin in a dimyristoylphosphatidylcholine bilayer membrane, *Biophys. J.* **1998**, *75*, 1603–1618.

[26] Wu, Y.; He, K.; Ludtke, S. J.; Huang, H. W., X-ray diffraction study of lipid bilayer membranes interacting with amphiphilic helical peptides: Diphytanoyl phosphatidylcholine with alamethicin at low concentrations, *Biophys. J.* **1995**, *68*, 2361–2369.

[27] Takahashi, S., Conformation of membrane fusion–active 20–residue peptides with or without lipid bilayers. Implication of α–helix formation for membrane fusion, *Biochemistry* **1990**, *29*, 6257–6264.

[28] Agawa, Y.; Lee, S.; Ono, S.; Aoyagi, H.; Ohno, M.; Taniguchi, T.; Anzai, K.; Kirino, Y., Interaction with phospholipid bilayers, ion channel formation, and antimicrobial activity of basic amphipathic α–helical model peptides of various chain lengths, *J. Biol. Chem.* **1991**, *266*, 20218–20222.

[29] Iwata, T.; Lee, S.; Oishi, O.; Aoyagi, H.; Ohno, M.; Anzai, K.; Kirino, Y.; Sugihara, G., Design and synthesis of amphipathic 3_{10}–helical peptides and their interactions with phospholipid bilayers and ion channel formation, *J. Biol. Chem.* **1994**, *269*, 4928–4933.

[30] Bechinger, B., Towards membrane protein design: pH–sensitive topology of histidine–containing polypeptides, *J. Mol. Biol.* **1996**, *263*, 768–775.

[31] Russell, C. J.; King, D. S.; Thorgeirsson, T. E.; Shin, Y. K., *De novo* design of a peptide which partitions between water and phospholipid bilayers as a monomeric α–helix, *Protein Eng.* **1998**, *11*, 539–547.

[32] Segrest, J. P.; Garber, D. W.; Brouillette, C. G.; Harvey, S. C.; Anantharamaiah, G. M., The amphipathic α–helix: A multifunctional structural motif in plasma apolipoproteins, *Adv. Prot. Chem.* **1994**, *45*, 303–369.

[33] Voglino, L.; McIntosh, T. J.; Simon, S. A., Modulation of the binding of signal peptides to lipid bilayers by dipoles near the hydrocarbon–water interface, *Biochemistry* **1998**, *37*, 12241–12252.

[34] Johnson, J. E.; Rao, N. M.; Hui, S. W.; Cornell, R. B., Conformation and lipid binding properties of four peptides derived from the membrane–binding domain of CFP: Phosphocholine cytidylifransferase, *Biochemistry* **1998**, *37*, 9509–9519.

[35] Chernomordik, L.; Chanturiya, A. N.; Suss-Toby, E.; Nora, E.; Zimmerberg, J., An amphipathic peptide from the C-terminal region of the human immunodeficiency virus envelope glycoprotein causes pore formation in membranes, *J. Viriol.* **1994**, *68*, 7115–7123.

[36] Ishiguro, R.; Kimura, N.; Takahashi, S., Orientation of fusion-active synthetic peptides in phospholipid bilayers: determination by Fourier transform infrared spectroscopy, *Biochemistry* **1993**, *32*, 9792–9797.

[37] Cajal, Y.; Rabanal, F.; Alsina, M. A.; Reig, F., A fluorescence and CD study on the interaction of synthetic lipophilic hepatitis B virus preS(120-145) peptide analogues with phopholipid vesicles, *Biopolymers* **1996**, *38*, 607–618.

[38] Bechinger, B.; Zasloff, M.; Opella, S. J., Structure and orientation of the antibiotic peptide magainin in membranes by solid–state nuclear magnetic resonance spectroscopy, *Protein Sci.* **1993**, *2*, 2077–2084.

[39] Bechinger, B.; Zasloff, M.; Opella, S. J., Structure and dynamics of the antibiotic peptide PGLa in membranes by solution and solid–state nuclear magnetic resonance spectroscopy, *Biophys. J.* **1998**, *74*, 981–987.

[40] Biggin, P. C.; Sansom, M. S., Simulation of voltage–dependent interactions of α–helical peptides with lipid bilayers, *Biophys. Chem.* **1996**, *60*, 99–110.

[41] Chung, L. A.; ; D., Lead J.; ; Degrado, W. F., Fluoresence studies of the secondary structure and orientation of a model ion channel peptide in phospholipid vesicles, *Biochemistry* **1992**, *31*, 6608–6616.

[42] O'Neil, K. T.; DeGrado, W. F., A thermodynamic scale for the helix–forming tendencies of the commonly occuring amino acids, *Science* **1990**, *250*, 646–651.

[43] Lazaridis, T.; Karplus, M., "New view" of protein folding reconciled with the old through multiple unfolding simulations, *Science* **1997**, *278*, 1928–1931.

[44] Sung, S. S.; Wu, X. W., Molecular dynamics simulations of synthetic peptide folding, *Proteins* **1996**, *25*, 202–214.

[45] Kohn, W. D.; Mant, C. T.; Hodges, R. S., α–helical protein assembly motifs, *J. Biol. Chem.* **1997**, *272*, 2583–2586.

[46] Skolnick, J.; Kolinski, A., Simulations of the folding of a globular protein, *Science* **1990**, *250*, 1121–1125.

[47] Sali, A.; Shakhnovich, E.; Karplus, M., How does a protein fold ?, *Nature* **1994**, *369*, 248–251.

[48] Dill, K. A.; Bromberg, S.; Yue, K.; Fiebig, K. M.; Yee, D. P.; Thomas, P. D.; Chan, H. S., Principles of protein folding — A perspective from simple exact models, *Prot. Sci.* **1995**, *4*, 561–602.

[49] Tirado-Rives, J.; Jorgensen, W. L., Molecular dynamics of the unfolding of an α–helical analogue of ribonuclease A S–peptide in water, *Biochemistry* **1991**, *30*, 3864–3871.

[50] Daggett, V.; Levitt, M., Molecular dynamics simluations of helix denaturation, *J. Mol. Biol.* **1992**, *223*, 1121–1138.

[51] Daggett, V.; Levitt, M., Protein unfolding pathways explored through molecular dynamics simulations, *J. Mol. Biol.* **1993**, *232*, 600–619.

[52] Duan, Y.; Kollman, P. A., Pathways to a protein folding intermediate observed in a 1–microsecond simulation in aqueous solution, *Science* **1998**, *282*, 740–744.

[53] Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E., Reversible peptide folding in solution by molecular dynamics simulation, *J. Mol. Biol.* **1998**, *280*, 925–932.

[54] Demchuk, E.; Bashford, D.; Case, D. A., Dynamics of a type IV turn in a linear peptide in aqueous solution, *Folding Des.* **1997**, *2*, 35–46.

[55] Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms, *J. Chem. Phys.* **1994**, *101*, 4177–4189.

[56] Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L., Explicit reversible integrators for extended systems dynamics, *Mol. Phys.* **1996**, *87*, 1117–1128.

[57] DeGrado, W. F.; Lear, J. D., Induction of peptide conformation at apolar/water interfaces. 1. A study with model peptides of defined hydrophobic periodicity, *J. Am. Chem. Soc.* **1985**, *107*, 7684–7689.

[58] Blondelle, S. E.; Ostreh, J. M.; Houghten, R. A.; Perez-Paya, E., Induced conformational states of amphipathic peptides in aqueous/lipid environments, *Biophys. J.* **1995**, *68*, 351–359.

[59] Kaiser, E.T.; Kezdy, F.J., Secondary structures of proteins and peptides in amphiphilic environments, *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 1137–1143.

[60] Beverung, C. J.; Radke, C. J.; Blanch, H. W., Adsorption dynamics of L–glutamic acid copolymers at a heptane/water interface, *Biophys. Chem.* **1998**, *70*, 121–132.

[61] Chipot, C.; Pohorille, A., Structure and dynamics of small peptides at aqueous interfaces. A multi–nanosecond molecular dynamics study., *J. Mol. Struct./Theochem — Proceedings of the fourth meeting of the World Association of Theoretically Oriented Chemists — WATOC '96* **1997**, *398/399*, 529–535.

[62] Zhong, Q.; Jiang, Q.; Moore, P. B.; Newns, D. M.; Klein, M. L., Molecular dynamics simulation of a synthetic ion channel, *Biophys. J.* **1998**, *74*, 3–10.

[63] Zhong, Q.; Moore, P. B.; Newns, D. M.; Klein, M. L., Molecular dynamics study of the LS$_3$ voltage–gated ion channel, *FEBS Lett.* **1998**, *427*, 267–270.

[64] Zhong, Q.; Hisslein, T.; Moore, P. B.; Newns, D. M.; Pattnaik, P.; Klein, M. L., The M2 channel of influenza A virus: A molecular dynamics study, *FEBS Lett.* **1998**, *434*, 265–271.

[65] Watts, A., Biophysics of the membrane interface, *Biochem. Soc. Trans.* **1995**, *23*, 959–65.

[66] White, S. H.; Wimley, W. C., Hydrophobic interactions of peptides with membrane interfaces, *Biochim. Biophys. Acta* **1998**, *1376*, 339–352.

[67] Rankin, S.; Watts, A.; Pinheiro, T. J., Electrostatic and hydrophobic contributions to the folding mechanism of apocytochrome *c* driven by the interaction with lipid, *Biochemistry* **1998**, *37*, 12588–12595.

[68] Owenson, B.; Wilson, M. A.; Pohorille, A. COSMOS — *A software package for COmputer Simulations of MOlecular Systems.* NASA — Ames Research Center, Moffett Field, CA 94035–1000, 1987.

[69] Verlet, L., Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules, *Phys. Rev.* **1967**, *159*, 98–103.

[70] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* **1983**, *79*, 926–935.

[71] Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., Optimized potential energy functions for liquid hydrocarbons, *J. Am. Chem. Soc.* **1984**, *106*, 6638–6646.

[72] Van Buuren, A. R.; Marrink, S. J.; Berendsen, H. J. C., A molecular dynamics study of the decane/water interface, *J. Phys. Chem.* **1993**, *97*, 9206–9212.

[73] Pohorille, A.; Wilson, M. A., Excess chemical potential of small solutes across water–membrane and water–hexane interfaces, *J. Chem. Phys.* **1996**, *104*, 3760–3773.
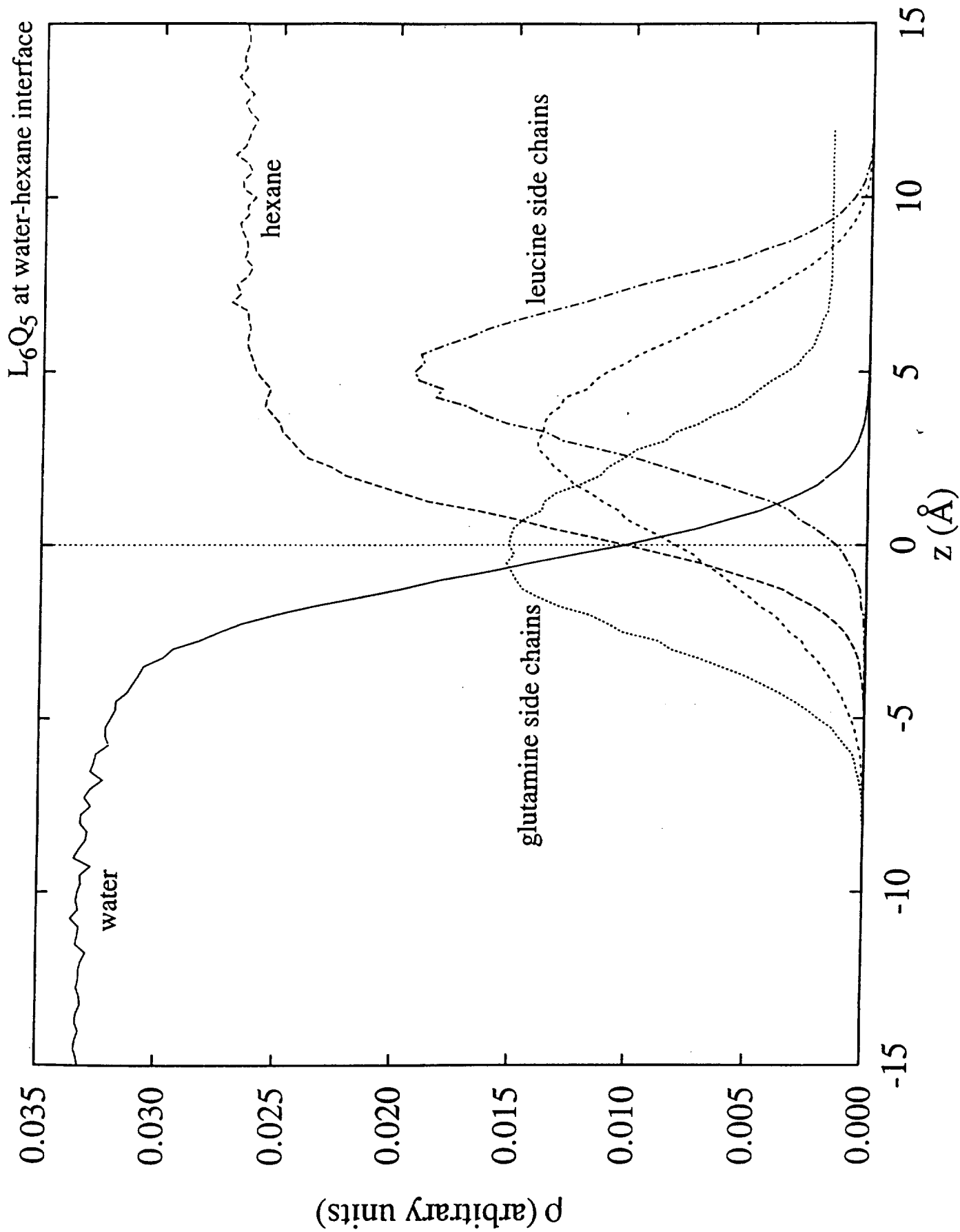
[74] Chipot, C.; Wilson, M. A.; Pohorille, A., Interactions of anesthetics with the water–hexane interface. A molecular dynamics study, *J. Phys. Chem. B* **1997**, *101*, 782–791.

[75] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr., K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. C.; Kollman, P. A., A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

[76] Wilson, M.; Pohorille, A., Molecular dynamics of a water–lipid bilayer interface, *J. Am. Chem. Soc.* **1994**, *116*, 1490–1501.

[77] Andrea, T. A.; Swope, W. C.; Andersen, H. C., The role of long–ranged forces in determining the structure and properties of liquid water, *J. Chem. Phys.* **1983**, *79*, 4576–4584.

[78] Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*, Clarendon Press: Oxford, 1987.

[79] Brooks III, C. L.; Pettitt, B. M.; Karplus, M., Structural and energetic effects of truncated long–ranged interactions in ionic and polar fluids, *J. Chem. Phys.* **1985**, *83*, 5897–5908.

[80] Ryckaert, J.; Cicotti, G.; Berendsen, H. J. C., Numerical integration of the Cartesian equations of motion for a system with constraints: Molecular dynamics of n–alkanes, *J. Comput. Phys.* **1977**, *23*, 327–341.

[81] Rudnick, J.; Gaspari, G., The shapes of random walks, *Science* **1987**, *237*, 384–389.

[82] Creighton, T. E. Conformational properties of polypeptide chains. in *Proteins, structures and molecular properties*. Freeman and Company, New York, 1993, pp. 171–199.

[83] Zimm, B. H.; Bragg, J. K., Theory of phase transition between helix and random coil in polypeptide chains, *J. Chem. Phys.* **1959**, *31*, 526–535.

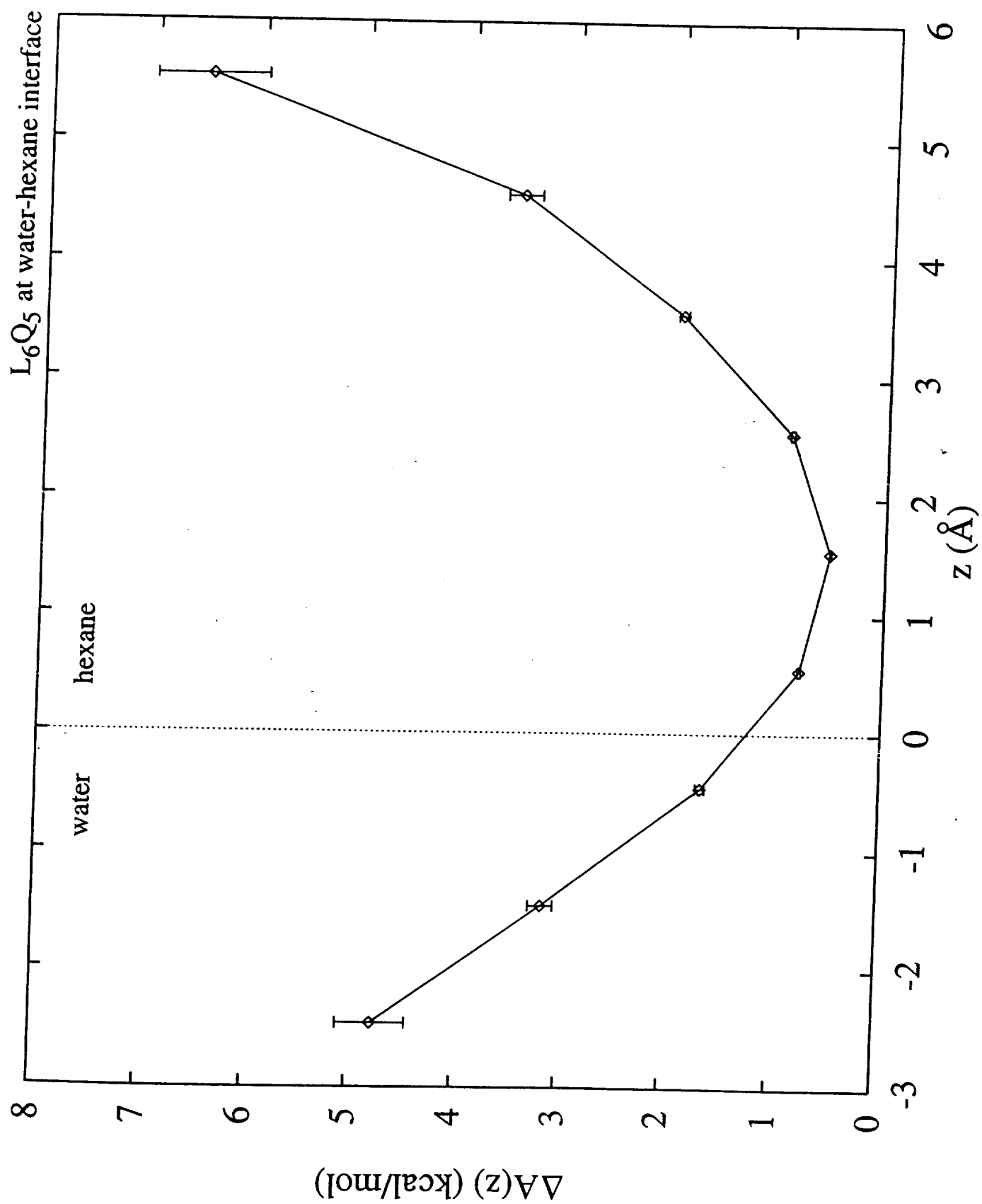[84] Poland, D.; Scheraga, H. A., *Theory of helix–coil transitions*, Academic Press: New York, 1970.

[85] Dobson, C. M.; Săli, A.; Karplus, M., Protein folding: A perspective from theory and experiment, *Angew. Chem. Int. Ed.* **1998**, *37*, 868–893.

[86] Dyson, H. J.; Rance, M.; Houghton, R. A.; Wright, P. E.; Lerner, R. A., Folding of immunogenic peptide fragments of proteins on water solution II. The nascent helix, *J. Mol. Biol.* **1988**, *201*, 201–217.

[87] Jähnig, F.; Bulow, R.; Baltz, T.; Overath, P., Secondary structure of the varian surface glycoproteins of trypanosomes, *FEBS Lett* **1987**, *221*, 37–42.

[88] Unson, C. G.; Gurzenda, E. M.; Iwasa, K.; Merrifield, R. B., Glucagon antagonists: contribution to binding and activity of the amino–terminal sequence 1-5, position 12 and the putative α–helical segment 19-27, *J. Biol. Chem.* **1989**, *264*, 789–794.

[89] Williams, R. W.; Weaver, J. L., Secondary structure of substance P bound to liposomes in organic solvents and in solution from Raman and CD spectroscopy, *J. Biol. Chem.* **1990**, *265*, 2505–2513.

[90] Van Den Hooven, H. W.; Doeland, C. C.; Van De Kamp, M.; Konings, R. N.; Hilberts, C. W.; Van De Ven, F. J., Three–dimensional structure of the lantibiotic nisin in the presence of membrane–mimetic micelles of dodecylphosphacholine and of sodium dodecylsulphate, *Eur. J. Biochem.* **1996**, *235*, 382–393.

[91] Rose, G. D.; Gierasch, L. M.; Smith, J. A., Turns in peptides and proteins, *Advan. Protein Chem.* **1985**, *37*, 1–109.

[92] Sundaralingam, M.; Sekharadu, Y. C., Water–inserted α–helical segments implicate reverse turns as folding intermediates, *Science* **1989**, *244*, 1333–1337.

[93] 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
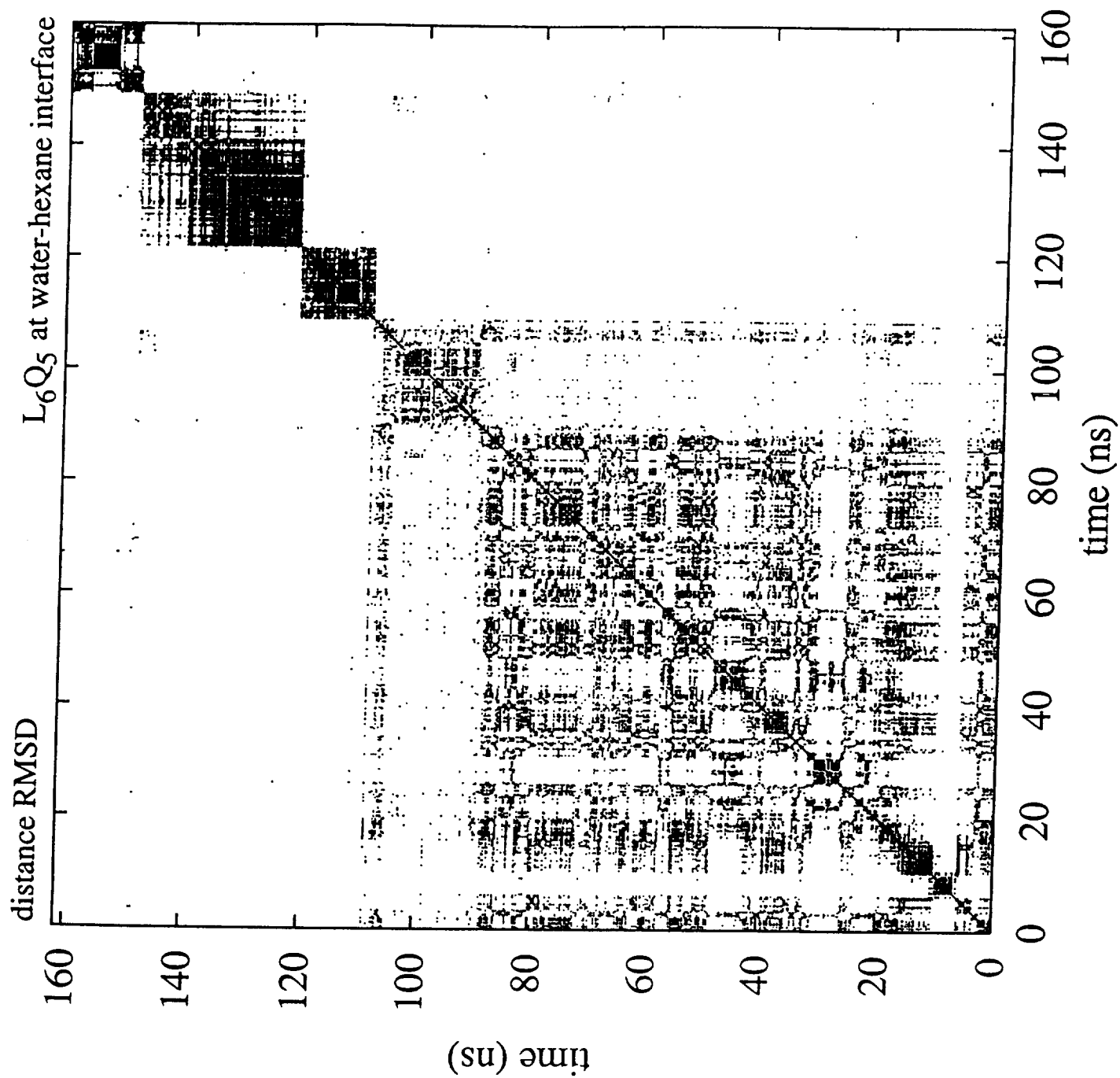
experimental data. in *Computer simulation of biomolecular systems: Theoretical and experimental applications*, Van Gunsteren, W. F.; Weiner, P. K., Eds. Escom, The Netherlands, 1997, pp. 83–96.

[94] Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free energy estimation: Umbrella sampling, *J. Comput. Phys.* **1977**, *23*, 187–199.

[95] Valleau, J. P.; Card, D. N., Monte Carlo estimation of the free energy by multistage sampling, *J. Chem. Phys.* **1972**, *57*, 5457–5462.

[96] Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The weighted histogram analysis method for free energy calculations on biomolecules. I. The method, *J. Comput. Chem.* **1992**, *13*, 1011–1021.

[97] Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A., Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields, *J. Am. Chem. Soc.* **1997**, *119*, 5908–5920.

[98] Stickle, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D., Hydrogen bonding in globular proteins, *J. Mol. Biol.* **1992**, *226*, 1143–1159.

[99] Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L., Short alanine-based peptides may form 3_{10}–helices in aqueous solution, *Nature* **1992**, *359*, 653–655.

[100] Fiori, W. R.; Miick, S. M.; Millhauser, G. L., Increasing sequence length favors α–helix over 3_{10}–helix in alanine–peptides: Evidence for a length–dependent structural transition, *Biochemistry* **1993**, *32*, 11957–11962.

[101] Millhauser, G. L., Views of helical peptides: A proposal for the position of 3_{10}–helix along the thermodynamic folding pathway, *Biochemistry* **1995**, *34*, 3873–3877.

[102] Soman, K. V.; Karimi, A.; Case, D. A., Unfolding of an α–helix in water, *Biopolymers* **1991**, *31*, 1351–1361.
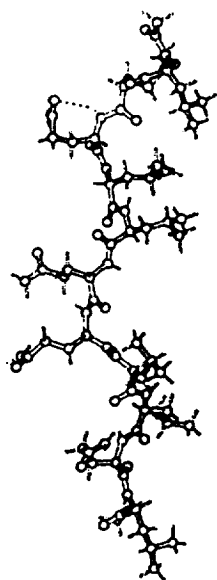
[103] Tobias, D. J.; Brooks III, C. L., Thermodynamics and mechanism of α-helix initiation in alanine and valine peptides, *Biochemistry* **1991**, *30*, 6059–6070.

[104] Tirado-Rives, J.; Maxwell, D. S.; Jorgensen, W. L., Molecular dynamics and Monte Carlo simulations favor the α-helical form for alanine-based peptides in water, *J. Am. Chem. Soc.* **1993**, *115*, 11590–11593.

[105] Millhauser, C. J.; Hanson, P.; Bolin, K. A.; van de Ven, F. J. M., Estimating the relative populations of 3_{10}-helix and α-helix in Ala-rich peptides: A hydrogen exchange in high field NMR study, *J. Mol. Biol.* **1997**, *267*, 963–974.

[106] Otoda, K.; Kitagawa, Y.; Kimura, S.; Imanishi, Y., Chain length dependent transition of 3_{10} to α-helix of BOC-(Ala-Aib)$_n$-OME, *Biopolymers* **1983**, *33*, 1337–1345.

[107] Basu, G.; Kuki, A., Evidence for a 3_{10}-helical conformation of an eight-residue peptide from ^{1}H-^{1}H rotating frame Overhauser studies, *Biopolymers* **1993**, *33*, 995–1000.

[108] Augspurger, J.D.; Bindra, V.A.; Scheraga, H.A.; Kuki, A., Helical stability of de novo designed α-aminoisobutyric acid-rich peptides at high temperatures, *Biochemistry* **1995**, *34*, 2566–2576.

[109] Zhang, Y. P.; Lewis, R. N. A. H.; Hodges, R. S.; McElhaney, R. N., Peptide models of helical hydrophobic transmembrane segments of membrane proteins. 1. Studies of the conformation, intrabilayer orientation and amide hydrogen exchangeability of Ac-K$_2$-(LA)$_{12}$-K$_2$-amide, *Biochemistry* **1995**, *34*, 2348–2361.

[110] Zhang, Y. P.; Lewis, R. N. A. H.; Hodges, R. S.; McElhaney, R. N., Peptide models of helical hydrophobic transmembrane segments of membrane proteins. 2. Differential scanning calorimetric and FTIR spectroscopic studies of the interaction of Ac-K$_2$-(LA)$_{12}$-K$_2$-amide with phosphatidylcholine bilayers, *Biochemistry* **1995**, *34*, 2362–2371.

[111] Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C., The helical hydrophobic moment: A measure of the amphipathy of a helix, *Nature* **1982**, *299*, 371–374.

[112] Wolfenden, R.; Andersson, L.; Cullins, P. M.; Southgate, C. C. B., Affinities of amino acid side chains for solvent water, *Biochemistry* **1981**, *20*, 849–855.

[113] Socci, N. D.; Onuchic, J. N.; Wolynes, P. G., Protein folding mechanisms and the multi-dimensional folding funnel, *Proteins: Struct. Funct. Genet.* **1998**, *32*, 136–158.

[114] Waltho, J. P.; Feher, V. A.; Merutka, G.; Dyson, H. J.; Wright, P. E., Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G– and H–helices of myoglobin, *Biochemistry* **1993**, *32*, 6337–6347.

[115] Bryson, J.W.; Desjarlais, J. R.; Handel, T.M.; DeGrado, W.F., From coiled coils to small globular proteins: Design of a native–like three–helix bundle, *Protein Sci.* **1998**, *7*, 1404–1414.
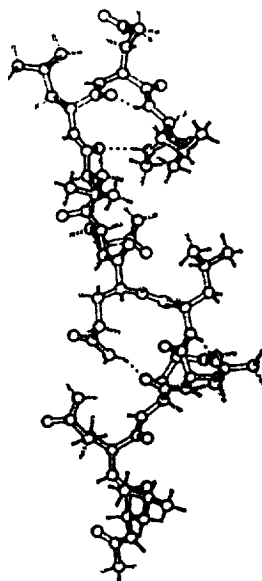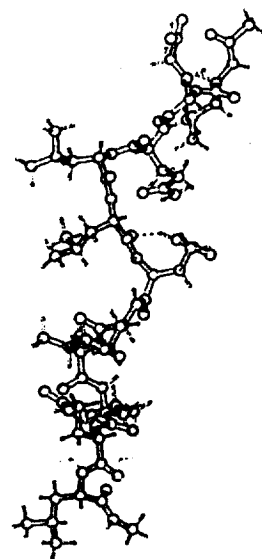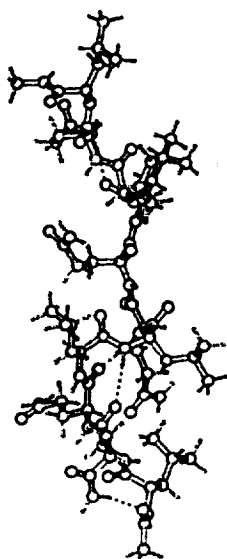
L_6Q_5 at water-hexane interface

hexane

leucine side chains

glutamine side chains

water

z (Å)

ρ (arbitrary units)

L_6Q_5 at water-hexane interface

distance RMSD

L_6Q_5 at water-hexane interface

time (ns)

time (ns)

t = 0 ns   t = 7.5 ns   t = 95 ns

t = 115 ns   t = 130 ns   t = 155 ns

residue 1

residue 2

residue 3

residue 4

residue 5

residue 6

L_6Q_3 at water-hexane interface

ψ (degrees)

φ (degrees)

180 120 60 0 -60 -120 -180

L₆Q₃ at water-hexane interface

L_6Q_5 at water-hexane interface

β

3_{10}

α_R

Asymmetry A_3

time (ns)

L_6Q_5 at water-hexane interface

α_R-helix

3_{10}-helix

time (ns)

distance RMSD (Å)

L_6Q_5 at water-hexane interface

time (ns)

N backbone-side chain H-bonds

L_6Q_5 at water-hexane interface