Insertion and Hairpin Formation of Membrane Proteins: A Monte Carlo Study

A. Baumgärtner
Institut für Festkörperforschung, Forschungszentrum, 52425 Jülich, Germany

ABSTRACT Some particular effects of a lipid membrane on the partitioning and the concomitant folding processes of model proteins have been investigated using Monte Carlo methods. It is observed that orientational order and lateral density fluctuations of the lipid matrix stabilize the orientation of helical proteins and induce a tendency of spontaneous formation of helical hairpins for helices longer than the width of the membrane. The lateral compression of the lipids on a hairpin leads to the extrusion of a loop at the *trans* side of the membrane. The stability of the hairpin can be increased by the design of appropriate groups of hydrophilic and hydrophobic residues at the extruded loop. It is shown that in the absence of lipids the orientation of proteins is not stable and the formation of hairpins is absent. Some analogies between the formation of helical hairpins in membranes and the formation of hairpins in polymer liquid crystals are discussed. The simulations indicate that the insertion process follows a well-defined pattern of kinetic steps.

INTRODUCTION

The partitioning of proteins into membranes and their subsequent folding are fundamental processes in biological cells (Engelman et al., 1986; Jacobs and White, 1989; Seelig and Ganz, 1991; Popot et al., 1994; White and Wimley, 1994; von Heijne, 1994a). Although in many cases proteins insert by taking advantage of a complex proteinaceous machinery (Wickner, 1979, 1994; Gierasch, 1989; Singer, 1990; Rapoport, 1992; Walter and Johnson, 1994), some, mostly short proteins are able to translocate spontaneously and *sec*-independently into membranes (Kuhn et al., 1990; Kilian et al., 1990; McKnight et al., 1991; Beschiachvili and Seelig, 1992; von Heijne, 1994b). Much progress has been made in elucidating the biochemical basis of spontaneous insertion of amphiphatic proteins into membranes (Eisenberg, 1984; Engelman et al., 1986; von Heijne, 1994a; White, 1994). However, various physical aspects of this process are not well understood, and several models for the transport of secretory proteins across or into membranes were proposed several years ago (von Heijne and Blomberg, 1979; Wickner, 1979; Engelman and Steitz, 1981; Jähnig, 1983).

One intriguing question is concerned with the role of the membrane, either passive or active, during the partitioning and folding processes of proteins (Jacobs and White, 1989; Popot and Engelman, 1990; Seelig and Ganz, 1991). Of course, the influence of the membrane on the translocation process can be expected to be quite complex, and therefore various aspects related to the thermodynamic state of the lipids, dipolar interfacial effects, membrane curvature (Baumgärtner and Skolnick, 1995; Baumgärtner, 1995), and others have to be addressed. Although it is conceivable that various effects do act in concert and would yield less effectivity if working separately, it is a useful strategy to investigate some effects of the membrane independently of each other.

In the present study we address the question of how the anisotropic orientational order of the lipids and their lateral density fluctuations affect the insertion of a protein. Introducing a simplified model membrane devoid of many details of lipid molecules, I simulate the insertion process of proteins of various chain lengths and compositions using the coarse-grained protein model of Milik and Skolnick (1993, 1995). The method of Monte Carlo simulation has been applied and preferred to molecular dynamics methods because of the expected slow dynamics of the translocation process.

In recent pioneering works Milik and Skolnick (1993, 1995) have studied the insertion of various proteins (M2δ, Melittin, Pf1, and fd). As in other simulations (Jähnig and Edholm, 1992; Sanders et al., 1991), the membrane was geometrically defined by a hydrophobic region, but virtually without lipids. The effect of the lipids on the protein had been taken into account by a mean orientational field (Milik and Skolnick, 1993). The basic and important result was that by using an appropriate hydropathy scale (Eisenberg, 1984; Engelman et al., 1986; Roseman, 1988; Jacobs and White, 1989), short membrane-spanning proteins successfully insert into a hydrophobic region.

The intention of the present work is to extend the previous studies by including lipid molecules during investigations. It will be shown that orientational order and lateral density fluctuations of the lipid matrix are important for the formation and stability of transmembrane helices and the formation of helical hairpins. Helical proteins with an axial contour length much larger than the width of the membrane exhibit a tendency to spontaneously form helical hairpins and a subsequent extrusion of coiled loops at the *trans* side of the membrane. The stability of the hairpin conformation

Received for publication 26 February 1996 and in final form 21 May 1996.
Address reprint requests to Dr. Artur Baumgärtner, IFF Forschungszentrum, D-52425 Julich, Germany. Tel.: 49-2461-61-4074; Fax: 49-2461-61-2620; E-mail: a.baumgaertner@kfa-juelich.de.
© 1996 by the Biophysical Society
0006-3495/96/09/1248/08 $2.00

can be enhanced by introducing hydrophilic side groups near the C-terminus and the N-terminus and in the extruded part of the protein.

MODELS OF POLYMER AND MEMBRANE

The membrane model

Based on our introductory arguments, we want to focus during our investigations on those effects on the protein insertion that are related to the orientational order and the lateral density fluctuations of the lipid matrix. Under this aspect we have constructed a model membrane that exerts essentially three constraints on a protein: 1) an entropic barrier during the partitioning process, 2) an anisotropic repulsive environment with lateral density fluctuations, and 3) a hydrophobic region. The simplest model fulfilling these requirements is a monolayer of hard parallel cylinders, representing the lipid chains, with orientations perpendicular to the x-y plane and confined to move stochastically in the x-y plane. Rotations of the cylinders with respect the z-direction are prohibited. The monolayer is an assembly of $M = 500$ cylindrical molecules ("lipids") confined in a unit box of area 15,900 Å^2 with periodic boundary conditions in the x-y domain. Each cylinder has a diameter of 4.9 Å and a length of 36 Å.

The area per chain is $a = 31.8$ Å^2, corresponding to a chain-chain spacing of $d = 6.05$ Å, according to $a = d^2\sqrt{3}/2$. This density corresponds approximately to the gel phase, where $d = 4.8$ Å ($a \approx 20$ Å^2), whereas in the fluid phase $d = 8.5$ Å ($a \approx 62.6$ Å^2) (see, e.g., Gennis, 1989). However, because the cross-sectional area of the chain is $a_o = 18.86$ Å^2, which is close to the area of an acyl chain in its all-*trans* state, but smaller than the average area per chain, $a = 31.8$ Å^2, a considerable amount of empty space between the chains is provided. This empty space has to be introduced to compensate for the impenetrability of the cylinders. Obviously, at closed packed densities, $a \approx a_o$, partitioning of a polymer into such a membrane is prohibited. So, the present discrepancy between the average area per chain and the cross-sectional area may be considered as a compromise to support protein insertion. However, it should be emphasized that insertion processes have been observed for higher density $a = 26.3$ Å^2 also, although on a much longer time scale. A systematic investigation of protein insertion at various densities of lipids is certainly of importance but is not within the scope of the present work.

The dynamics of the lipids is achieved by randomly displacing the lipids in the x-y plane. Each move is rejected if it leads to a violation of excluded volume conditions among the cylinders or between cylinders and the protein.

The protein model

In calculating the conformational properties of a polymer, it is sometimes convenient to consider the virtual bonds from one residue to the neighboring one. This has been suggested by Flory (1969) and has been applied to proteins by several authors (Miller and Flory, 1966; Gregoret and Cohen, 1990; Milik and Skolnick, 1993). The C-α representation is used, and the average virtual bond is 3.8 Å. In the model of Gregoret and Cohen (1987) the amino acids are represented by two or three spheres. A different Monte Carlo model including side chains has been proposed very recently by Sung (1994, 1995).

More recently, Milik and Skolnick (1993, 1995) have used this concept and have developed a coarse-grained protein model suitable for fast dynamic Monte Carlo methods. In the Milik-Skolnick model a single sphere of diameter σ_P is used for every amino acid and placed at the junctions of the virtual bonds ("stick-and-ball chain"). In the present study we have used the Milik-Skolnick model with some minor modifications.

The total energy of the protein consists of various contributions,

$$U = U_{\text{bond}} + U_{\text{angle}} + U_{\text{torsion}} + U_{\text{steric}} + U_S + U_H. \quad (1)$$

Parameters for U have been estimated by various groups (van Gunsteren and Karplus, 1982; Brooks et al., 1983; Mayo et al., 1990).

The vibrational energy of the bonds along the chain backbone is

$$U_{\text{bond}} = \sum_{\text{bonds}} \epsilon_b (b/b_0 - 1)^2, \quad (2)$$

where $\epsilon_b = 3.6$ kcal/mol and $b_0 = 3.8$ Å. We used an upper and lower bound for the extension of the spring such that U_{bond} ($b > b_{\max}$) = $U_{\text{bond}}(b < b_{\min}) = \infty$, where $b_{\min} = 3.4$ Å, $b_{\max} = 4.2$ Å.

The bending energy between neighboring bonds is

$$U_{\text{angles}} = \sum_{\text{angles}} \epsilon_\theta (\cos\theta - \cos\theta_0)^2, \quad (3)$$

where $\epsilon_\theta = 2.0$ kcal/mol and $\theta_0 = 89.5$.

The torsional energy is

$$U_{\text{torsion}} = \sum_{i,\, i+1} \epsilon_\phi \{1 - \cos[n(\phi - \phi_0)]\}, \quad (4)$$

where ϕ is the dihedral angle between two successive planes spanned by the bonds $\mathbf{r}_{ij}$, $\mathbf{r}_{jk}$, and $\mathbf{r}_{kl}$, n is the periodicity, $\phi_0 = 52.1$ is the equilibrium angle, and $\epsilon_\phi = 1.5$ kcal/mol.

The fourth term in Eq. 1 takes into account the steric interactions among residues and between residues and lipids:

$$U_{\text{steric}} = \sum_{i=1}^{N} \sum_{j=1}^{M} V(r_{ij}), \quad (5)$$

where V is a hard sphere potential

$$V(r_{ij}) = \begin{cases} 0, & \text{for } |\mathbf{r}_i - \mathbf{r}_j| > \sigma \\ \infty, & \text{for } |\mathbf{r}_i - \mathbf{r}_j| < \sigma \end{cases}, \quad (6)$$

where $\sigma \equiv \sigma_P$ for residue-residue contact, and $\sigma \equiv (\sigma_P + \sigma_L)/2$ for residue-lipid contact. The hard sphere diameter of a residue is $\sigma P = 3.0$ Å, and the diameter of a lipid is $\sigma_L = 4.9$ Å. N is the number of residues of the protein, and M is the number of lipids in the membrane. In general, the effective diameter of a side-chain group depends slightly on the type of amino acid (Gregoret and Cohen, 1990). This has been neglected in the present study.

The fifth term in Eq. 1 is due to the hydropathies of the amino acid side chains and the hydrophilicity of the polypeptide backbone of the protein. The corresponding total energy of transport of the protein from a polar aqueous environment to the hydrophobic core of the membrane is given by

$$U_S = \sum_{j=1}^{N} [s_j + c_j] g(z_j), \quad (7)$$

where s_j and c_j are the average energy of transport of a side-chain group and a peptide group, respectively. The s_j depends on the type of amino acid side chain. Estimates of various hydropathy scales have been reported and discussed by several authors (Eisenberg, 1984; Engelman et al., 1986; Roseman, 1988; Jacobs and White, 1989). In the present work we have used the parameters of the Roseman scale (Roseman, 1988). The parameter c_j represents the hydrophilicity of the peptide group and takes into account the loss of hydrogen bond energy of a backbone group with water during the insertion into the hydrophobic region of the membrane. They are assumed to be independent of j, and $c_j \equiv c = 4$ kcal/mol. Because in the present study water molecules are not explicitly taken into account, one must introduce an interfacial profile function that describes the change of fraction of water with distance $z_j - z_o$ from the interfacial plane of the membrane located at z_o. This interfacial function is assumed to vary

exponentially:

$$g(z_j) = \frac{1}{1 + \exp[(z_j - z_o)/\lambda]}, \quad (8)$$

and the decay length λ has been assumed to be on the order of the interfacial regime between hydrocarbon core and the polar environment, which is on the order of $\lambda = 5$ Å. Similar profile functions have been used in other studies (Jähnig and Edholm, 1992; Milik and Skolnick, 1993).

The last term in Eq. 1 takes into account the formation of internal hydrogen bonds. The corresponding total energy is given by

$$U_H = \sum_{j=1}^{N} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2}[h_j\, g(z_j) + h_i\, g(z_i)]V_H(r_{ij}). \quad (9)$$

This energy compensates for the loss of water-peptide hydrogen bonds in the polar solvent, represented by the term $\sim c_j g(z_j)$ in Eq. 7, during the partitioning of the chain into the hydrophobic core of the membrane.

The energy parameter $(h_i + h_j)/2$ is the average hydrogen bond energy of the backbone groups i and j. In general, the average hydrogen bond energy of each residue in the chain depends weakly on the type of amino acid side group because of different steric interactions. However, in the present work we neglect these differences and assume for all residues the same parameter $h_j \equiv h = -6.6$ kcal/mol. This choice yields the experimentally observed value of about 6 kcal/mol.

In general, the second sum of Eq. 9 has to be performed over $n = N$ neighboring residues. However, because of conformational restrictions in the helical state, the main contributions are related to the third and the fourth neighbors along the chain. Here we restrict the second summation in Eq. 9 to these two neighbors. The factor $1/2n$, with $n = 2$ in the case of third and fourth nearest neighbors, is a normalization factor providing an average energy per residue j, which is necessary to compare average hydrogen bond energies to the average segregation energies s_j and c_j as used in Eq. 7.

The phenomenological hydrogen bond potential $V_H(r_{ij})$ has to be accommodated to the C-α representation of the present protein model and has been proposed (Milik and Skolnick, 1993) as

$$V_H(r_{ij}) = \frac{1}{1 + [(|\mathbf{r}_i - \mathbf{r}_j| - a_{ij})/\kappa]^4}. \quad (10)$$

The potential has a maximum, $V_H(a_{ij}) = 1.0$, at the optimal distance $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j| = a_{ij}$ between two residues i and j. The decay length κ of the potential is assumed to $\kappa = 0.1\, a_{ij}$. Because the axial shift per residue in an α-helix is 1.5 Å, the optimal distance a_{ij} between the fourth and third nearest neighbors is approximately 6 Å and 4.5 Å, respectively. More accurate values are $a_3 = 5.04$ Å and $a_4 = 6.3$ Å (Barlow and Thornton, 1988), which have been used in the present work.

The Monte Carlo method (Binder, 1984) is used to calculate the dynamic and equilibrium properties of the model system. The dynamics of the lipids and the protein are achieved by randomly displacing their constituents. Each move is accepted if

$$\exp(-\Delta U/RT) > \eta, \quad (11)$$

where $0 < \eta < 1$ is a random number, and ΔU is the difference between the old and new total energy of the system. Otherwise the old configuration of the system is counted as the new one. Throughout the simulations we used a temperature of $T = 305$K, and hence $RT = 0.606$ kcal/mol. The displacements of polymer and lipids are performed at equal rates. One Monte Carlo step is $M + N$ attempted moves and is defined as one time unit.

The initial configuration of the polymer is an equilibrated random coil, i.e., the polymer is placed far from the membrane surface, say 200 Å, and equilibrated according to Eq. 1 for about 7000 Monte Carlo steps, which is comparable to the typical configurational relaxation time on the order of N^2. The equilibrium state is reached if the mean square fluctuation of the radius of gyration remains constant for larger times. After equilibration the chain is placed in proximity to the membrane surface such that the closest distance of any of the residues is not smaller than 10 Å from the surface. The subsequent partitioning process is then monitored.

RESULTS AND DISCUSSION

Previous studies on protein insertion (Milik and Skolnick, 1993) were mainly concerned with testing appropriate hydropathy scales applied to particular single membrane-spanning proteins, such as M2δ, melittin, Pf1, and others. In the present work we want to focus mainly on the partitioning and subsequent folding processes. To avoid specific yet unknown effects due to the heterogeneity of the amino acid sequence on partitioning and folding, we rather study proteins with a homogeneous sequence of the type X_6Y_n, where Y_n denotes a sequence of n identical amino acids and X_6 denotes a sequence of six amino acids serving as a hydrophilic anchor at the surface on the *cis* side of the membrane. For the present purpose the anchor sequence must not be very specific, and we have chosen as an appropriate candidate a modified version of the C-terminus sequence of melittin (see, e.g., Dempsey, 1990), X = Gly—Gln—Arg—Lys—Arg—Lys—. For the sequence Y_n we have used leucine, which is one of the most hydrophobic residues, $s = -3.02$ kcal/mol, according to the Roseman scale (Roseman, 1988).

Insertion of a short protein

The simulation of a short protein of the type X_6 Leu_{28} has been performed to test the ability of the present model protein to insert into the lipid matrix and to exhibit the expected coil-helix transition during the insertion process. The length of the hydrophobic part, Leu_{28}, has been chosen to match the length of the helix with the width of the membrane.

The average coordinate $\langle z_n \rangle$ of the nth residue perpendicular to the membrane surface is presented in Fig. 1. Two

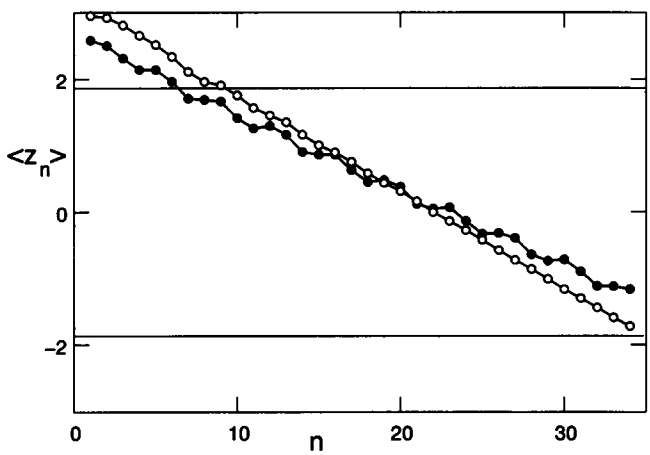

FIGURE 1 Average $\langle z_n \rangle$ coordinate of residues n along the chain for the short protein X_6 Leu_{28} (○). ●, Case where lipid-peptide interactions are ignored. $\langle z_n \rangle$ is given in nanometers.

snapshots of the protein from the side and top are depicted in Figs. 2 and 3, respectively. These results conform to the expected helical structure (Milik and Skolnick, 1993), with a pitch of about 5.4 Å and a rise per residue of about 1.5 Å.

The orientation of the helix is perpendicular the membrane surface, and the fluctuations about this orientation are small compared to the case where steric interactions between protein and lipids are ignored. This is demonstrated in Fig. 4 by the angular distribution function P (cos θ) of the helical axis with respect to the membrane surface. Without steric interactions the distribution is much broader than in the lipid-protein case, which indicates that the orientational order of the lipids imposes a significant constraint on the orientation of the helical axis. It should be noted that this difference cannot be clearly deduced from the average coordinates $\langle z_n \rangle$, as shown in Fig. 1, where the two curves corresponding to the two cases do not differ too much from each other.

It should be noted that the hydrophilic strength of the anchor sequence X_6 is sufficiently strong to pin this terminus of the protein at the interface, and hence specific Coulomb interactions between terminus group and interface are not required.

Insertion of a long polypeptide

Because a short protein cannot react upon the finite width of the membrane, it is of interest to examine the behavior of much longer hydrophobic sequences. Then several conflicting scenarios could be envisaged: either no partitioning at all, or protrusion of the C-terminus into the hydrophilic region beyond the *trans* membrane side, or bending back of the helix. The results of the simulations of the protein X_6 Leu_{56} are shown in Figs. 5–8.

The curves in Fig. 5 represent a typical time series of conformations during the insertion process. The coordinates $z_n(t)$ at four different times t indicate the successive formation of a helical hairpin. The following successive kinetic steps take place. The middle part of the chain starts at first to penetrate into the lipid matrix, and the α-helix develops simultaneously from both ends of the chain toward the middle. These two processes promote each other and lead to the formation of a stable helical turn located about the middle of the chain. It is important to note that the same pattern of kinetic steps is observed, starting with various different initial random chain conformations outside the membrane.

The average equilibrium values of $\langle z_n \rangle$ are presented in Fig. 6 and denoted by the open circles. Snapshots of the helical hairpin are depicted in Figs. 7 and 8 as views from the side and top, respectively.

The importance of the lipid matrix with respect to the formation of a helical hairpin is stressed by comparison with the behavior of the same protein when the steric interactions between lipids and protein are ignored. In this case the data of $\langle z_n \rangle$, denoted by the full circles in Fig. 6, rather provide

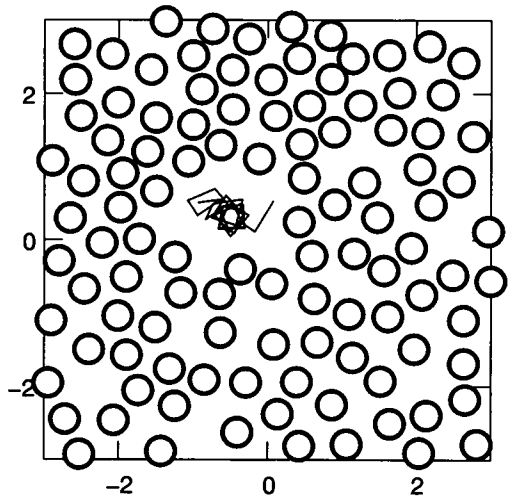

FIGURE 3 Snapshot of X_6 Leu_{28} from top, including the lipid molecules. The circles represent the lipid molecules from the top.

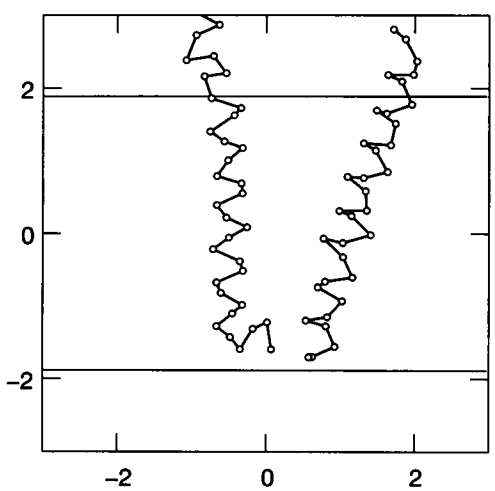

FIGURE 2 Snapshot of X_6 Leu_{28}. ○, Positions of the C_α atoms.

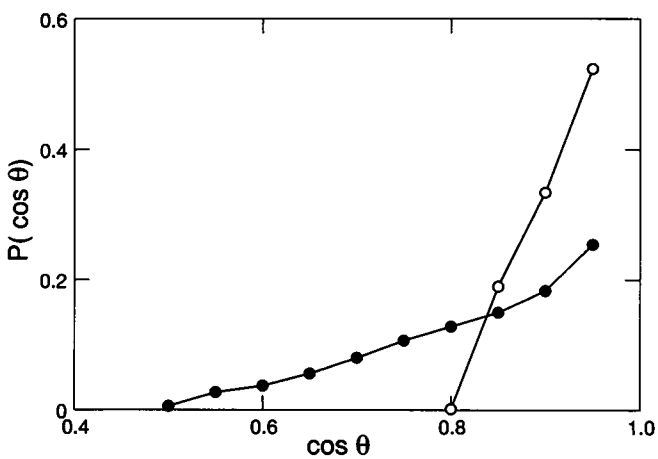

FIGURE 4 Probability distribution P(cos) of the orientation of the helical axis of X_6 Leu_{28}. ●, Case where lipid-protein interactions are ignored.

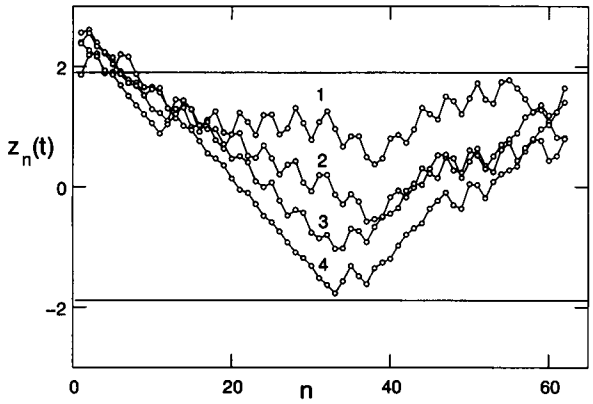

FIGURE 5 Insertion process of a long protein X_6 Leu_{56} characterized by successive ensembles of the $z_n(t)$ coordinates of the n residues along the chain at different times t. The numbers at each curve denote the Monte Carlo time in units of 6000 time steps.

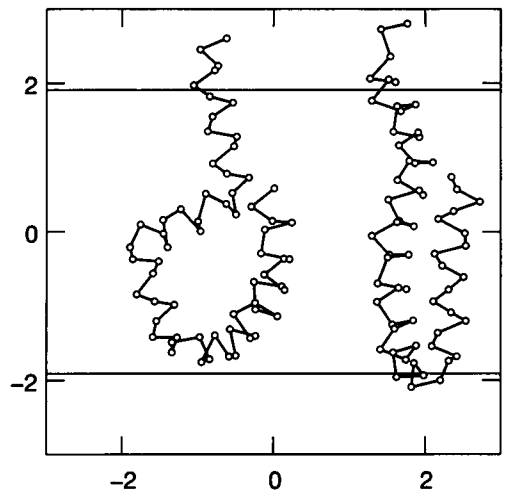

FIGURE 7 Snapshot of X_6 Leu_{56}. ○, Positions of the C_α atoms.

evidence of a distorted helix extending almost parallel between the two membrane surfaces.

From these observations one can conclude that if the hydrophobic sequence prevents the protein from extension into the hydrophilic environment at the *trans* side of the membrane, then the helical molecule accommodates to the anisotropic lipid matrix by the formation of a helical hairpin. It is conceivable that the concomitant loss of energy by breaking hydrogen bonds at the turn of the hairpin must be compensated by the gain of "steric" free energy by rendering semirigid helical segments parallel to the orientation of the lipids, quite similar to lyotropic liquid crystals (Onsager, 1949; Flory, 1984).

Formation of hairpin and loop

However, the exact location of the turn along the helical axis is, in principle, not restricted. It may fluctuate slowly around some average location and could lead to some conformational instabilities. One way to stabilize the helical hairpin would be to introduce proline (von Heijne, 1991) instead of leucine, approximately near residue $n = 28$ to favor a kink at this position. Another possibility encountered in many membrane proteins is the extrusion of the helical turn at the *trans* side of the membrane into the polar environment. This could be promoted by replacing the residues in the neighborhood of the turn by hydrophilic side groups. Of course, the replacements cannot be arbitrary and could even prevent protein insertion. The turning of size and strength of this hydrophilic loop is a delicate task generally related to the prediction of sequences and structures of membrane proteins (Engelman et al., 1986; von Heijne, 1994a) and will not be pursued in the present work. Rather, we will demonstrate at a particular protein the robustness of

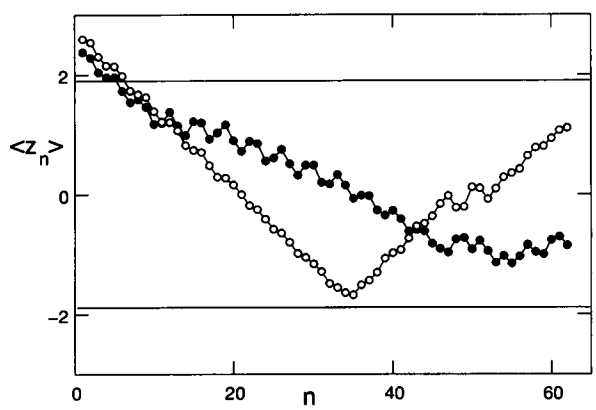

FIGURE 6 Average $\langle z_n \rangle$ coordinate of residues n along the chain for X_6 Leu_{56} (○). ●, Case where lipid-protein interactions are ignored.

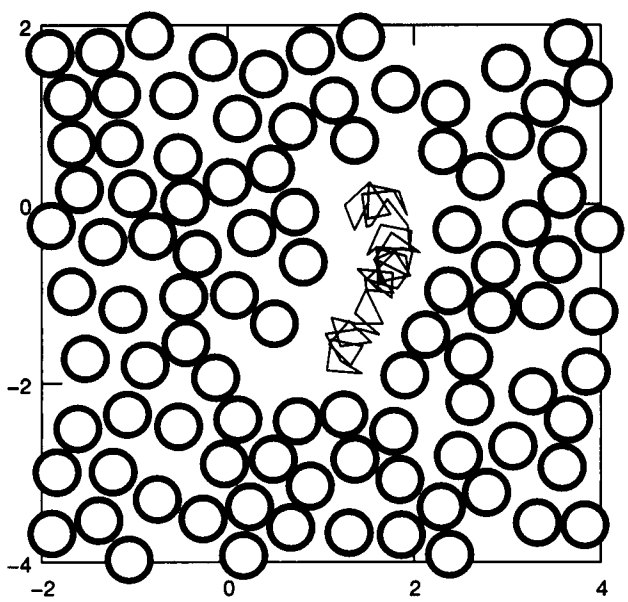

FIGURE 8 Snapshot of X_6 Leu_{56} from top, including the lipid molecules. The circles represent the lipid molecules from the top.

the translocation process, including the formation of a helical hairpin.

The formation and extrusion of such a loop and its stabilizing effect on a helical hairpin can be studied by simulating a protein of the type X_6 Leu_{24} Gly_{20} Leu_{24} X_6. The two sequences X_6 at both ends serve as anchors at the *cis* side of the membrane, as in the case of the protein M13 (Kuhn et al., 1990). The group Gly_{20}, which has a hydrophobicity of $s \approx 0$ kcal/mol (Roseman, 1988), can be expected to form a loop extending into the polar region at the *trans* side of the membrane. The results of the simulations are presented in Figs. 9–12.

The process of formation of the helical hairpin and the subsequent extrusion of a loop are depicted in Fig. 9 by means of $z_n(t)$ at various times t. After the adhesion to the *cis* side of the membrane surface (time = 1), the two hydrophobic Leu sequences start to penetrate into the lipid matrix (time = 3) and successfully pull the Gly sequence into the hydrophobic region (time = 9). At that stage two helical hairpins, at about $n \approx 25$ and $n \approx 60$, had formed. Each length fluctuates at the expense of the other, until one of the hairpins dominates (times = 13, 26, 30 in Fig. 9). This leads to the localization of the turn at the Gly sequence. The turn becomes finally extruded (time = 50 in Fig. 9), forming a coiled loop outside the membrane. Because various simulations starting with different chain conformations outside of the membrane exhibit similar time evolutions of partitioning and formation of helical hairpins, it seems very likely that the insertion process in general follows a certain pattern of successive kinetic steps, as described above.

Typical snapshots of the final helical hairpin are portrayed in Figs. 10 and 11 from side and top, respectively. The data of $\langle z_n \rangle$ in the final equilibrium state are shown in Fig. 12 (denoted by the open circles), and they indicate a stable helical hairpin with an orientation parallel to the lipid molecules. For comparison, the corresponding data of the same protein but without steric interactions with the lipids are included in Fig. 12, denoted by the full circles. The statistical error in the former case is approximately 4 Å, whereas in the latter case it is considerably larger because of orientation fluctuations, and is probably on the order of 10 Å. According to the average coordinates there is no evidence of a helical hairpin, but rather a distorted bend helix pervading the "empty" hydrophobic region. The latter case clearly demonstrates the importance of the lipid matrix for

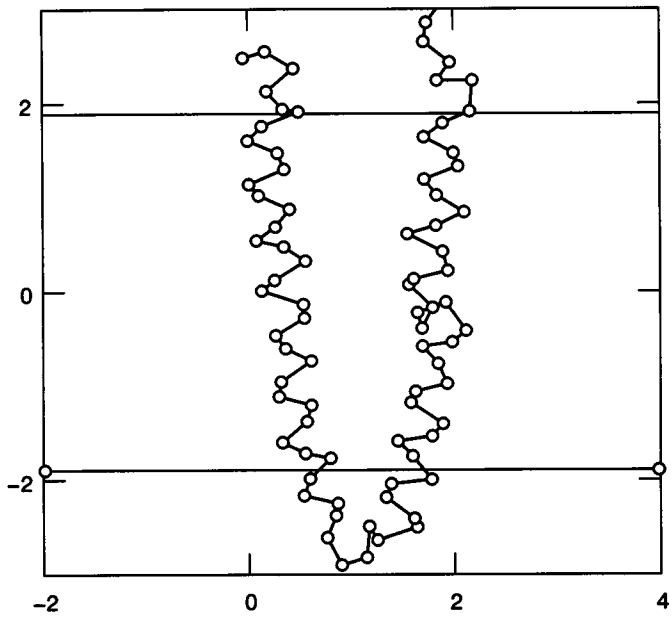

FIGURE 10 Snapshot of X_6 Leu_{24} Gly_{20} Leu_{24} X_6. ○, Positions of the C_α atoms.

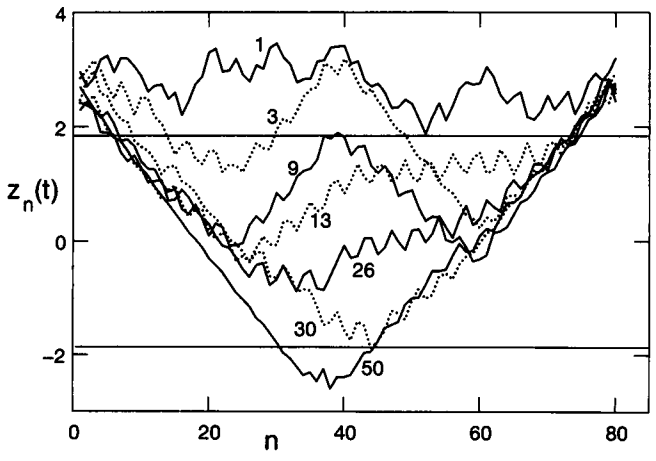

FIGURE 9 Insertion process and formation of a helical hairpin of the protein X_6 Leu_{24} Gly_{20} Leu_{24} X_6 characterized by the $z_n(t)$ coordinates of the n residues along the chain at subsequent times t. The numbers at each curve denote the Monte Carlo time in units of 1000 time steps.

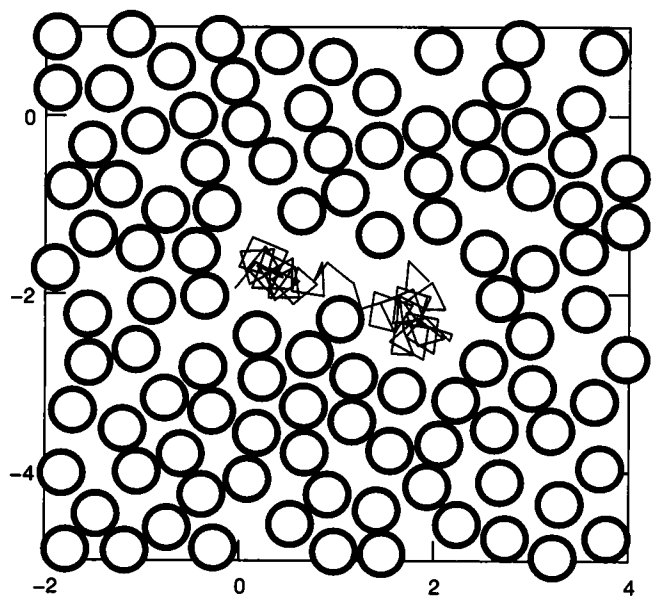

FIGURE 11 Snapshot of X_6 Leu_{24} Gly_{20} Leu_{24} X_6 from the top, including the lipid molecules. The circles represent the lipid molecules from the top.

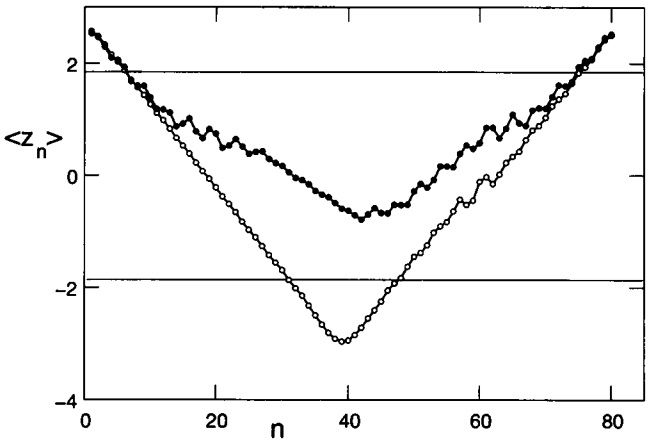

FIGURE 12 Average z_n coordinate of residues n along the chain for X_6 Leu_{24} Gly_{20} Leu_{24} X_6 (○). ●, Case where lipid-protein interactions are ignored.

the formation and the stability of a helical hairpin. Without the anisotropic steric interaction between the helix and the lipids, the formation of a hairpin would be virtually impossible. The lateral pressure imposed by the fluctuating lipids on the growing helices after the partitioning into the membrane promotes the formation of a hairpin and leads finally to the extrusion of an extramembranous loop.

It is interesting to note that the formation of a hairpin as a special type of conformation of a macromolecule is not restricted to helical membrane proteins only, but can be encountered in polymer liquid crystals as well (de Gennes, 1982; Warner et al., 1985). It is well known that periodic polymers consisting of a repeating unit of short, rodlike, nematogenic molecules separated by flexible spacers may exhibit in a nematic environment a considerable amount of hairpin defects. In this case there are two main reasons for the formation of hairpins: the flexibility of the spacers, which allows spontaneous reversal of the chain direction, and the nematogenic character of the rodlike groups forcing, in concert with the anisotropic environment, the chain to meander in accord with the nematic director. Obviously, hairpins are an important contribution to the conformational entropy of the chain. Comparing liquid crystal polymers with helical membrane proteins, even with regard to multiple membrane-spanning proteins, it is tempting to suggest some analogies by identifying the semirigid membrane-spanning helices and their connecting coiled loops with the repeating units of nematogenic polymers, which are the rodlike molecules separated by flexible spacers. How far these analogies could be used to improve our knowledge of protein insertion is still an open question and may be considered in the future.

SUMMARY AND CONCLUSIONS

In the present work the effects of orientational order and lateral density fluctuations on the insertion and the formation of a hairpin of some special proteins have been investigated using Monte Carlo simulations. It is found that orientational order and lateral density fluctuations of the lipids stabilize the orientation of helical proteins and induce the spontaneous formation of helical hairpins. The lateral compression of the lipids on a hairpin leads to the extrusion of a loop at the *trans* side of the membrane. The stability of the hairpin can be increased by the design of appropriate groups of hydrophilic and hydrophobic residues at the extruded loop. It is shown that in the case of the absence of lipids, the orientation of a protein is not stable and the formation of a hairpin is not observed. The simulations indicate that the insertion process follows a certain pattern of successive kinetic steps. Although several simplifications with respect the intermolecular interactions have been assumed, the essential conclusions based upon the present results seem to be promising enough to justify further investigations along this line using more elaborate model membranes and proteins.

Of course, several other effects observed during protein insertion have been deliberately neglected in the present model simulations. Among others, the nonclassical hydrophobic effect (Seelig and Ganz, 1991; Beschiaschvili and Seelig, 1992), which is supposed to be related to compensating contributions from the desolvation of an amphiphilic molecule and the increased hydration of the bilayer interface during the partitioning process, is of importance for a complete understanding of the thermodynamic basis of protein insertion. Other interesting effects, related to Coulomb interactions among residues and with surface charges, which may be of importance for targeting to and transport across membranes (Kuhn et al., 1990), might be of relevance as well.

REFERENCES

Barlow, D. J., and J. M. Thornton. 1988. Helix geometry in proteins. *J. Mol. Biol.* 201:601–619.

Baumgärtner, A. 1995. Asymmetric partitioning of a polymer into a curved membrane. *J. Chem. Phys.* 103:10669–10674.

Baumgärtner, A., and J. Skolnick. 1995. Spontaneous translocation of a polymer across a curved membrane. *Phys. Rev. Lett.* 74:2142–2145.

Beschiaschvili, G., and J. Seelig. 1992. Peptide binding to lipid bilayer. Nonclassical hydrophobic effect and membrane-induced pK shifts. *Biochemistry.* 31:10044–10053.

Binder, K., editor. 1984. Applications of the Monte Carlo Method in Statistical Physics. Springer, Heidelberg.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.

de Gennes, P. G. 1982. Mechanical properties of nematic polymers. *In* Polymer Liquid Crystals. A. Ciferri, W. R. Krigbaum, and R. B. Meyer, editors. Academic Press, New York.

Dempsey, C. E. 1990. The actions of melittin on membranes. *Biochim. Biophys. Acta.* 1031:143–161.

Eisenberg, D. 1984. Three-dimensional structure of membrane and surface proteins. *Annu. Rev. Biochem.* 53:595–623.

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

Engelman, D. M., T. A. Steitz, and A. Goldman. 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* 15:321–353.

Flory, P. J. 1969. Statistical Mechanics of Chain Molecules. John Wiley and Sons, New York.

Flory, P. J. 1984. Molecular theory of liquid crystals. *Adv. Polym. Sci.* 59:1–36.

Gennis, R. B. 1989. Biomembranes: Molecular Structure and Function. Springer, New York.

Gierasch, L. M. 1989. Signal sequences. *Biochemistry.* 28:923–930.

Gregoret, L. M., and F. E. Cohen. 1990. Novel method for the rapid evaluation of packing in protein structures. *J. Mol. Biol.* 211:959–974.

Jacobs, R. E., and S. H. White. 1989. The nature of the hydrophobic binding of small peptides at a bilayer interface: implications for the insertion of transbilayer helices. *Biochemistry.* 28:3421–3437.

Jähnig, F. 1983. Thermodynamics and kinetics of protein incorporation into membranes. *Proc. Natl. Acad. Sci. USA.* 80:3691–3695.

Jähnig, F., and O. Edholm. 1992. Modeling the structure of bacteriorhodopsin. A molecular dynamics study. *J. Mol. Biol.* 226:837–850.

Kilian, J. A., R. C. Keller, M. Struyve, A. de Kroon, J. Tommassen, and B. de Kruiff. 1990. Tryptophan fluorescence study on the interaction of a signal peptide of the *Escherichia coli* outer membrane protein PhoE with model membranes. *Biochemistry.* 29:8131–8137.

Kuhn, A., J. Rohrer, and A. Gallusser. 1990. Bacteriophages M13 and Pf3 tells us how proteins insert into the membrane. *J. Struct. Biol.* 104: 38–43.

Mayo, S. L., B. D. Olafson, and W. A. Goddard. 1990. DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* 94: 8897–8909.

McKnight, C. J., M. Rafalski, and L. M. Gierasch. 1991. Fluorescence analysis of tryptophan-containing variants of the LamB signal sequence upon insertion into a lipid bilayer. *Biochemistry.* 30:6241–6246.

Milik, M., and J. Skolnick. 1993. Insertion of peptide chains into lipid membranes: an off-lattice Monte Carlo dynamics model. *Proteins Struct. Funct. Genet.* 15:10–25.

Milik, M., and J. Skolnick. 1995. A Monte Carlo model of fd and Pf1 proteins in lipid membranes. *Biophys. J.* 69:1382–1386.

Miller, W. G., and P. J. Flory. 1966. Dimension of polypeptide chains in heliogenic solvents. *J. Mol. Biol.* 15:298–314.

Onsager, L. 1949. The effects of shape on the interaction of colloidal particles. *Ann. N.Y. Acad. Sci.* 51:627–659.

Popot, J. P., and D. M. Engelman. 1990. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry.* 29:4031–4037.

Popot, J. P., C. Vitry, and A. Atteia. 1994. Folding and assembly of integral membrane protein. *In* Membrane Protein Structure. S. H. White, editor. Oxford University Press, New York. 41–96.

Rapoport, T. A. 1992. Transport of proteins across the endoplasmatic reticulum membrane. *Science.* 258:932–936.

Roseman, M. A. 1988. Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. *J. Mol. Biol.* 200:513–522.

Sanders, J. C., N. van Nuland, O. Edholm, and M. A. Hemminga. 1991. Conformation and aggregation of M13 coat protein studied by molecular dynamics. *Biophys. Chem.* 41:193–202.

Seelig, J., and P. Ganz. 1991. Nonclassical hydrophobic effect in membrane binding equilibria. *Biochemistry.* 30:9354–9359.

Singer, S. J. 1990. The structure and insertion of integral proteins in membranes. *Annu. Rev. Cell. Biol.* 6:247–296.

Sung, S.-S. 1994. Helix folding simulations with various initial conformations. *Biophys. J.* 66:1796–1803.

Sung, S.-S. 1995. Folding simulations of Alanine-based peptides with lysine residues. *Biophys. J.* 68:826–834.

van Gunsteren, W. F., and M. Karplus. 1982. Effects of constraints on the dynamics of macromolecules. *Macromolecules.* 15:1528–1544.

von Heijne, G. 1991. Proline kinks in transmembrane α-helices. *J. Mol. Biol.* 218:499–503.

von Heijne, G. 1994a. Membrane proteins: from sequence to structure. *Annu. Rev. Biophys. Biomol. Struct.* 23:167–192.

von Heijne, G. 1994b. *Sec*-independent protein insertion into the inner *E. coli* membrane: a phenomenon in search of an explanation. *FEBS Lett.* 346:69–72.

von Heijne, G., and C. Blomberg. 1979. Trans-membrane translocation of proteins. The direct transfer model. *Eur. J. Biochem.* 97:175–181.

Walter, P., and A. E. Johnson. 1994. Signal sequence recognition and protein targeting to the endoplasmatic reticulum membrane. *Annu. Rev. Cell. Biol.* 10:87–119.

Warner, M., J. M. F. Gunn, and A. Baumgärtner. 1985. Rod to coil transitions in nematic polymers. *J. Phys. A. Math. Gen.* 18:3007–3026.

White, S. H. 1994. Hydropathy plots and the prediction of membrane protein topology. *In* Membrane Protein Structure. S. H. White, editor. Oxford University Press, New York. 97–124.

White, S. H., and W. C. Wimley. 1994. Peptides in bilayers: structural and thermodynamic basis for partitioning and folding. *Curr. Opin. Struct. Biol.* 4:79–86.

Wickner, W. T. 1979. Assembly of proteins into membranes: the membrane trigger hypothesis. *Annu. Rev. Biochem.* 48:23–45.

Wickner, W. T. 1994. How ATP drives proteins across membranes. *Science.* 266:1197–1198.