16

Monte Carlo Models of Spontaneous Insertion of Peptides into Lipid Membranes

JEFFREY SKOLNICK AND MARIUSZ MILIK

Introduction

The export of proteins from one compartment to another requires that they be transported across cellular membranes. In addition, membrane proteins must somehow insert into the lipid bilayer after they have been synthesized. Because of their biological importance, the problems of protein insertion into membranes as well as the mechanism(s) of transport across membranes have been extensively studied (Singer, 1990). Experimental data strongly suggest that in vivo most large membrane proteins do not spontaneously insert into membranes (Singer, 1990). Rather, a complex cellular machinery is used. While the basic elements of this machinery are known for many systems, the mechanism and the source of the translocation energy are still unclear (Gierash, 1989; Rapoport, 1992; 1991). On the other hand, when the translocation machinery is blocked or absent, many in vitro experiments provide evidence that some short proteins can insert into phospholipid vesicles (Maduke and Roise, 1993) or into membranes (Wolfe et al, 1985). These experimental results suggest that long and short proteins may translocate by different mechanisms in the cell (von Heijne, 1994) or that the translocation machinery is used to catalyze what is fundamentally a spontaneous insertion process (Jacobs and White, 1989); this may be particularly true for the case of long, amphiphilic sequences. Thus, a number of experiments suggest that spontaneous insertion can occur in membrane–protein systems and that spontaneous insertion can be explained on the basis of the thermodynamics of these systems. However, the detailed mechanism of such spontaneous insertion is not fully understood, nor is the panoply of interactions that drive the insertion process fully characterized. In this review, we describe recent Monte Carlo simulations that are designed to address these important questions.

Biological Membranes
K. Merz, Jr. and B. Roux, Editors

Membrane–protein systems are inherently complicated multicomponent systems having a very large number of degrees of freedom. Therefore, they are very difficult to describe from either a thermodynamic or microscopic viewpoint. These systems exhibit a very complex balance between the enthalpy and entropy of the protein insertion process. The change in entropy associated with peptide insertion can be substantial. In addition to the structural changes of the lipid accompanying protein insertion, the free energy of the water will also change. One also has to deal with changes in hydrogen bond energy, as well as van der Waals and electrostatic interactions. Thus, if one wishes to examine the mechanism of insertion of a peptide from water into a lipid, in order to make the problem computationally tractable, simplified models must be studied. Their advantage is that they can span the time scale of the insertion process; on the other hand, a number of the fine atomic details may not be accounted for. Whether or not such details provide the dominant interactions will determine the validity of the simulations described in this chapter. However, comparison with the full atom simulations of Roux and coworkers (see the chapter by Roux and Woolf in the present volume) indicates that in a number of respects these simplified models have captured the essence of the physics.

The simulations presented here represent different levels of simplification. We begin with a very schematic lattice model of a membrane bound peptide chain, then turn to a more realistic off-lattice model exhibiting a more elaborate membrane structure. Finally, we describe models where the structure of the membrane is represented by set of schematic lipid molecules, but the possibility of conformational transitions in the peptide is ignored. Depending on the level of detail, different questions can be answered by these simulations. Overall, though, the objective of these simulations is to provide insights into the general mechanism of peptide insertion, to explore whether the orientation and conformation of membrane peptides can be predicted, and to examine some possible roles the membrane may play in the peptide insertion process.

Lattice Model of Peptide Insertion

Lattice models, while schematic and somewhat artificial, offer a number of advantages that are particularly important in the case of modeling of large and complicated biomolecular systems (Godzik et al, 1993). The possibility of precalculating the set of Monte Carlo moves that modify the chain configuration and the use of simplified interaction potentials can result in a substantial reduction in simulation time (sometimes by up to three orders of magnitude) in comparison with the analogous off-lattice model. Thus, certain problems only become computationally tractable when a lattice representation is used. However, the advantage of computational speed may be counterbalanced by the imposition of some artificial, geometrical constraints onto the modeled system. Obviously, some questions related to fine structural details cannot be addressed. Fortunately, by going to a sufficiently high resolution lattice, lattice

simulations have proven to be very useful for modeling the basic thermodynamic and kinetic properties of systems such as globular proteins (Kolinski and Skolnick, 1994).

Description of the Model

The first lattice model used to investigate the mechanism of spontaneous insertion of peptides into lipid membranes is due to Milik and Skolnick (Milik and Skolnick, 1992). They describe the model peptide molecule as a sequence of diamond-lattice vectors, where three successive vectors represent one peptide residue (amino acid). Thus, the model schematically depicts the backbone NH, C^{α} and carbonyl groups. Here, the peptide chain can form three regular kinds of secondary structure: right and left handed helices (with four residues per turn) and an expanded zigzag state, which is a low resolution model of β structure. In order to simplify the calculation, the lipid is modeled as a structureless, effective medium which simply occupies a region of space, and the details associated with the existence of lipid molecules, which in reality form the lipid bilayer, are ignored. The bilayer only implicitly enters by its effect on the values of the hydrophobic and hydrogen bond potentials which depend on the z coordinate of the residue. Additional details follow below.

The Potential

The potential used in the lattice model contains four parts: a term, depicting how the peptide residues interact with each other and with the virtual lipid phase; a torsion potential for the peptide chain that represents the intrinsic propensity of residues to adopt a given type of secondary structure; a hydrogen bond potential; and an ordering potential which is designed to imitate steric interactions of the peptide chain with lipid molecules.

The interaction between peptide residues is represented by a mean-force Lennard-Jones (L-J) type potential (8–4), centered on the C^{α} carbons. The repulsive part of L-J potential is used when residues (e.g., hydrophils with hydrophobs) do not have any attractive region:

$$v_{LJ} = 4\varepsilon[(\sigma/r)^8 - (\sigma/r)^4]. \tag{1}$$

The parameters ε are different for different residue types. The σ parameter is constant and is defined such that for residues having an attractive region, V_{LJ} has a minimum at the distance equal to the length of one diamond lattice vector; i.e., $\sigma = \sqrt{3}/\sqrt[4]{2}$. The magnitude of the L-J interaction also depends on the z coordinates of both residues (depending on whether the residue is hydrophobic, hydrophilic, or lipophilic, pair interactions between residues are different in the water and in the hydrocarbon phase).

The energy of transport of a residue from the water into the hydrocarbon–membrane environment is represented by a one-body hydrophobic potential which depends on its type and its z coordinate (this mimics the effect of the

Table 1a. Values of Hydrophobic Potential for Model Residues

Symbol of residue	E_{hph} [kT]
a	4.0
b	0.0
c	−1.5
d	−1.5
e	−4.5
f	−4.5

Table 1b. Values of ε Parameters of the Two-Body Interaction for Pairs of Residues Inside of the Model Membrane

	a	*b*	*c*	*d*	*e*	*f*
a	0.0	0.2	1.0	0.5*	2.0*	2.0*
b		0.2	0.0	0.5*	0.5*	0.5*
c			1.5	0.5*	1.5	0.5*
d				0.5*	0.5*	0.5*
e					1.5	0.5*
f						0.5*

* Values with asterisks denote a repulsive interaction between the residues.

virtual lipid membrane). Within the bilayer region, hydrophobic residues have a lower energy, whereas hydrophilic residues have a higher energy. The converse holds true in the aqueous region of space.

In order to represent the variety of natural amino acids, six types of peptide residues are used, which differ by values of the L-J ε parameters, as well as by the one-body hydrophobic potential values. The values of the parameters for the different residue types used in these simulations are presented in Tables 1a and 1b.

The hydrogen bond potential is independent of residue type and may be depicted by:

$$E_{Hb} = \begin{cases} \varepsilon_{Hb}, \text{ when } r_{ij} = dist_1 \\ 0, \text{ when } dist_1 < r_{ij} < dist_2 \\ 0, \text{ when } r_{ij} > dist_2 \text{ and atoms are in the lipid phase} \\ \varepsilon_{HbW}, \text{ when } r_{ij} > dist_2, \text{ and atoms are in the water phase} \end{cases} \quad (2)$$

where: r_{ij} is the distance between the nitrogen or carbonyl atom of the residue i and the appropriate atom of residue j; $dist_1$ is the length of the sum of two diamond lattice vectors; $dist_2$ is the length of the sum of three diamond lattice

vectors in the trans conformation; $\epsilon_{Hb} = -1.5\ kT$ is the energy of intramolecular hydrogen bond interaction; and $\epsilon_{HbW} = -0.5\ kT$ is the energy of the hydrogen bond of a peptide with water.

Secondary structure propensities are controlled by a torsion potential. In the case of the diamond lattice model, a very simple potential has been used which discriminates against left-handed helices. Both right-handed helical and extended conformations are isoenergetic. Furthermore, in the bilayer region, the domination of helical structures is due to better saturation of hydrogen bonds.

The membrane phase, composed of lipid hydrocarbon chains, is ordered and anisotropic. One might expect that a fragment of the peptide chain which is buried in the membrane will locally disrupt this order and that the magnitude of the disruption will depend on the orientation of the fragment with respect to the bilayer. In the lattice model, this property of the membrane system has been addressed by introducing an orientational energy term of the form:

$$E_{ord} = \varepsilon_{ord} \sin^2(\theta) \tag{3}$$

where: ε_{ord} is a coefficient and θ is the angle between the end-to-end vector of a peptide fragment and the main lipid axis (which lies along the z coordinate). The value of the ε_{ord} parameter was small in comparison with the other potential parameters used in the present model and assumes values from 0.15 to 0.3 kT.

Results

The model has been used to simulate the spontaneous insertion process of two types of transmembrane structures whose sequences are given in Table 2. The first sequence consists of two identical hydrophobic, transmembrane helices connected by a hydrophilic linker. The second is more complicated and has been chosen to emulate the composition of the N-terminus of secreted proteins (having an uncleaved signal sequence). The composition of this peptide can be schematically depicted as: (charged fragment) – (hydrophobic fragment) – (hydrophilic linker) – (amphipathic fragment).

In all of the computational experiments presented here, the model peptide chain starts from a random conformation in the water phase. All the runs have been repeated 15 times, with different seeds for the random number generator. In the case of hydrophobic helices with a hydrophilic linker, the insertion process starts with adsorption of the chain on the lipid interface. This adsorbed,

Table 2. Sequences Used in the Lattice Model Simulations

Sequence number	*N* terminus	First transbilayer fragment	Linker	Second transbilayer fragment	*C* terminus
1	bb	ccddccddccddccddccdd	aaaaaa	ccddccddccddccddccdd	bb
2	aa	eeffeeffeeffeeffeeff	bbbbbb	ccddccddccddccddccdd	aa

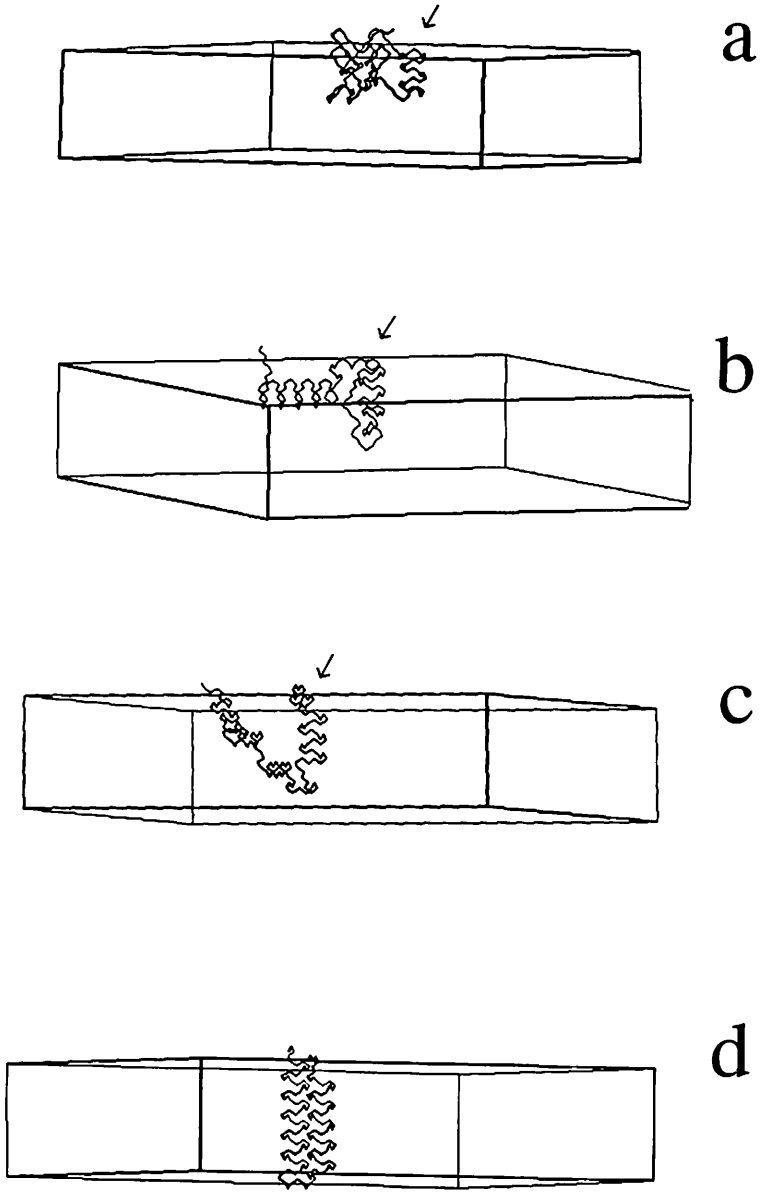

Figure 1. Snapshots of the insertion process of the diamond lattice chain into the model membrane. The arrows point to the N-terminus of the chain.

unordered structure represents a deep and broad local free energy minimum, and the chains spend about half of the simulation time in this minimum. The transition into the global free energy minimum for this sequence (a transmembrane hairpin) is associated with a large energetic barrier (about 20 kT) caused by the need to disrupt peptide-water hydrogen bonds. However, in all simulation runs, both chain ends eventually cross the membrane and result in the formation of a stable, transmembrane structure.

The second sequence has been prepared to illustrate Engelman and Steitz's helical hairpin hypothesis (Engelman and Steitz, 1981), which was later extended by Jacobs and White (1989). According to this hypothesis, fragments of proteins transported through the membrane first preassemble at the membrane interface, and then they are transported through the membrane as helical blocks. Our simulations, in which the sequence of the chain emulates an initial fragment of secreted protein, confirms this mechanism of transport. Figure 1 displays critical points of this process. First, the chain is adsorbed onto the membrane surface (Figure 1a). Then, the amphipathic fragment forms an adsorbed helix whose main axis is parallel to the interface, and the hydrophobic fragment forms a helical structure whose main axis is perpendicular to the interface. Both fragments are connected by a long, disordered linker (Figure 1b). Gradually, the transmembrane structure grows and pulls the amphipathic fragment through the membrane (Figure 1c), and finally, the antiparallel, transmembrane, helical structure forms (Figure 1d). The diamond lattice simulations indicate that the insertion process of peptide chains into the lipid membrane can be explained on a schematic level as arising from the effects of hydrophobic and hydrogen bond interactions.

Off-Lattice Model of Peptide Insertion

As mentioned above, lattice models may impose some artificial, geometrical constraints which could affect the polypeptide structure and dynamics. Therefore, an off-lattice extension of the model depicted above is a natural next step (Milik and Skolnick, 1993). Additionally, this offers the possibility of more realistically representing both the peptide chain and the model of the membrane.

Description of the Model

The protein is represented by a chain of balls with centers at the C^{α} carbon positions. The distance between adjacent balls equals the mean distance between neighboring C^{α} carbons in protein structures from the Brookhaven Protein Data Bank. The radii of these balls are taken from a statistical analysis of a database of protein structures done by Gregoret and Cohen (1990). The Monte Carlo Dynamics method is used to calculate the equilibrium properties of this model. The model system diffuses in conformational space by a set of discrete micromodifications of the chain conformation. The probability of acceptance for a

newly generated state depends on the energies of old and new states, according to the asymmetric Metropolis scheme (Binder, 1984). The set of elementary chain micro-modifications consists of spike moves for the ends and internal residues and long range sliding moves. On the level of secondary structure, uniform (i.e., residue independent) helical propensities are assumed for all the simulations and for all residues.

Interaction Scheme

Proceeding in a similar fashion to the lattice model described above, the hydrophobic effect exerted by the membrane is introduced as a z coordinate dependent effective potential which represents the different environments associated with the water, the lipid, and the lipid-water interface. Thus, the hydrophobic interaction energy for the residue of interest is not only a function of the amino acid type but also of its z coordinate in the Monte Carlo box. In contrast to the schematic lattice model described above, in this case the hydropathy scale is derived from the experimental data of Jacobs and White (1987) for the interaction of tripeptides with model lipid bilayers and from Roseman's partition data for model compounds in octanol/water systems (Roseman, 1988). Additionally, the self-solvation effect is used to scale the thermodynamic data, according to Roseman's proposition (Roseman, 1988).

The interface of a lipid membrane contains mostly lipid head groups. It has very specific thermodynamic properties, which are different from both the hydrocarbon and water phases. Some experimental data indicate that certain amino acids interact specifically with this region (Jacobs and White, 1987). This is suggestive of the importance of the interfacial region in the first stage of the insertion process, the adsorption of a peptide on the membrane interface. Therefore, the interface region is treated more elaborately here than it has been in the previous lattice model.

Conceptually, the process of transport of a protein fragment from water into the hydrocarbon phase can be divided into two main stages: adsorption on the interface and transport from the interface into the hydrocarbon phase. According to Jacobs and White (1989), the free energy of peptide adsorption onto the membrane interface is represented by:

$$\Delta G_{w \to if} = \Delta G_{ifh} + \Delta G_{imm} \tag{4}$$

where ΔG_{ifh} is the free energy of partially burying the protein fragment in the lipid phase (hydrophobicity), and ΔG_{imm} is the change of free energy due to reduction of external degrees of freedom of the peptide chain.

The burial energy, ΔG_{ifh}, is approximated in the present model as a linear function of the total accessible surface area of a residue, according to equation:

$$\Delta G_{ifh} = f C_s A_s \tag{5}$$

where C_s is a solvation parameter, taken from the literature (Jacobs and White, 1989), that equals -22.0 cal/(Mol Å^2), and A_s is the total accessible surface

area of the peptide in an extended state (Lesser et al, 1987). f represents the fractional area of a residue that is buried in the adsorbed state; for the simulations described here, $f = 0.56$.

The change of free energy related to the reduction of the external degrees of freedom, ΔG_{imm}, is more difficult to estimate (Jahnig, 1983); it is omitted in the present model because the entropic part of the ΔG_{imm} is at least partially considered in the Monte Carlo scheme.

The free energy of transport of a protein fragment in an extended conformation from the interface into the hydrocarbon phase is defined analogously to the Jacobs and White definition (Jacobs and White, 1989):

$$\Delta G_{ext} = (1 - f)C_s A_{Gly} + \Delta G_{bb} + \Delta G_{sc} \tag{6}$$

where: A_{gly} denotes the accessible surface area of a glycine residue (used as a representation of backbone surface area). ΔG_{bb} is the free energy of transfer of non-hydrogen-bonded backbone polar groups. In the present model, it is assigned a value of 4.1 kcal/Mol (Milik and Skolnick, 1993). ΔG_{sc} is the free energy of side chain transfer from water into a hydrocarbon phase; it has been corrected for the self-solvation effect. The values for the naturally occurring amino acids are obtained from Roseman (Roseman, 1988).

Finally, the transport of a helical protein fragment from the membrane interface into the hydrocarbon phase is related to the following free energy change:

$$\Delta G_{hlx} = \Delta G_{ext} + \Delta G_{Hb} + \Delta G_{lost} \tag{7}$$

where ΔG_{Hb} is the free energy of hydrogen bond formation, set here to -3.8 kcal/Mol. ΔG_{lost} represents the decrease of side chain hydrophobicity due to the change of the accessible surface area of side chains during formation of a helical conformation; the values of this parameter are taken from Roseman (Roseman, 1988). Table 3 contains a compendium of the calculated values of ΔG_{ext} and ΔG_{hlx} for the naturally occurring amino acids.

Finally, all protein fragments (in either helical or extended conformations), when transported into the lipid membrane, disrupt the packing of lipid chains, and therefore they change the ordering of the hydrocarbon phase. As in the case of the lattice model (Equation 3), the change of free energy associated with this effect may be calculated from:

$$\Delta G_{lip} = C_{ord} \sin^2(\theta) \tag{8}$$

where: C_{ord} is a coefficient with values ranging from 0.05 to 0.15 kcal/Mol, and θ is the angle between the end-to-end vector of a polypeptide fragment comprised of four residues and the normal to the membrane surface (z axis).

Results

All simulations described below commence with a random chain conformation located somewhere in the aqueous phase. Most simulations have run for 5×10^6 Monte Carlo steps, where one Monte Carlo step is defined as the time in

Table 3. Values (in kcal/mol) of Free Energy of Transport from the Interface to the Interior of the Lipid Membrane for Amino Acids in Nonhelical (ΔG_{ext}) and Helical (ΔG_{hlx}) Conformations

Residue	ΔG_{ext}	ΔG_{hlx}
Ala	2.58	−0.82
Cys	1.25	−2.05
Asp	4.82	1.82
Glu	5.03	2.53
Phe	0.01	−3.09
Gly	3.25	−0.55
His	4.34	1.54
Ile	0.23	−3.17
Lys	5.71	2.31
Leu	0.23	−2.97
Met	1.58	−1.92
Asn	5.52	2.62
Pro	−1.10	−0.60
Gln	5.37	2.27
Arg	7.14	3.64
Ser	3.35	−0.05
Thr	2.83	−0.57
Val	1.07	−2.33
Trp	0.39	−2.81
Tyr	4.23	1.63

which the algorithm tried on average to move every residue of the model chain. Information about the system is stored every 2500 Monte Carlo-steps into a trajectory file. This file is then used for the analysis of the behavior of the model system during the simulation.

Sequences of known, natural membrane bounded peptides and bacteriophage coat proteins have been used in the computational experiments described here. The sequences are presented in the one-letter amino acid notation (Bechinger et al, 1991); bold fonts indicate hydrophobic amino acids.

MAGAININ2

Magainin2 is a member of a family of peptides found in frog skin. Magainins have some antibacterial activity, and, according to NMR experiments, they are predominantly helical in membranes. The sequence of magainin2 is: GIGKFLH SAKKFGKAFVGEIMNS.

Figure 2a presents the final conformation of the chain. In the model, the magainin2 chain is not transported into the hydrocarbon phase; rather, it forms a highly stable, helical structure, which is adsorbed on the interface. This final configuration is found in all runs on the magainin2 sequence.

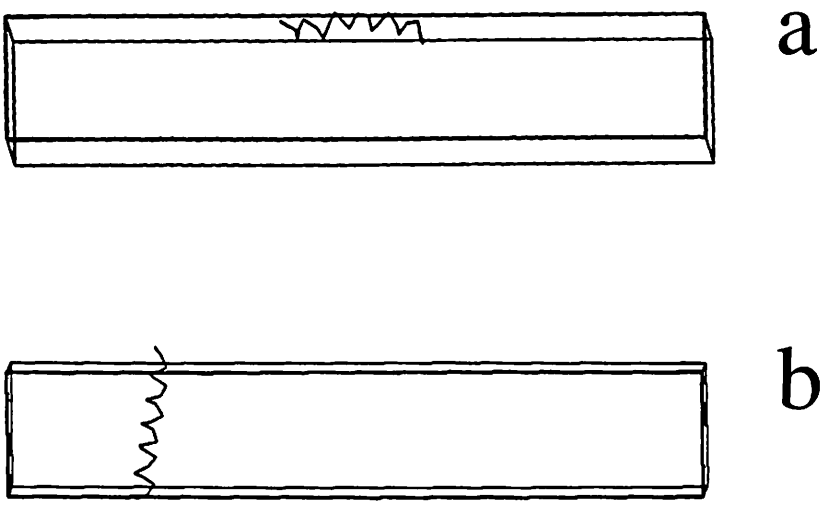

Figure 2. Predicted orientation and structure of (a) magainin2, and (b) M2δ peptides.

Figure 3a contains additional information about this simulation, which confirms that magainin2 adsorbs onto the membrane interface with one face of the helix formed by hydrophobic residues. The mean value of the z coordinates for the chain residues are presented in Figure 3a as a function of the residue number in the sequence. The bars on the figure denote the values of the mean absolute deviations of the data. Based on Figure 3a, we predict that the magainin2 chain forms a stable, helical structure adsorbed on the membrane interface. As indicated by the relatively small values of the mean deviations, vertical mobility is negligible.

M2δ PEPTIDE

The M2δ peptide is a peptide that forms transmembrane helical structures and has the sequence (Bechinger et al, 1991): EKMSTAISVLLAQAVFLLLTSQR. The hydrophobic residues form clusters in the middle of the M2δ sequence, thereby rendering its hydrophobic pattern similar to signal peptide sequences. On the other hand, in magainin2 the hydrophobic and hydrophilic residues are mixed together. The contrasting behavior of the two sequences lends support to the supposition that differences in hydrophobicity pattern, rather than overall sequence hydrophobicity, decide whether a given peptide forms a transmembrane structure.

Simulations on the M2δ peptide start from the same initial configurations and physico-chemical conditions as is the case for magainin 2. Figure 2b presents the final, transmembrane structure, which is stable during the last 80% of the

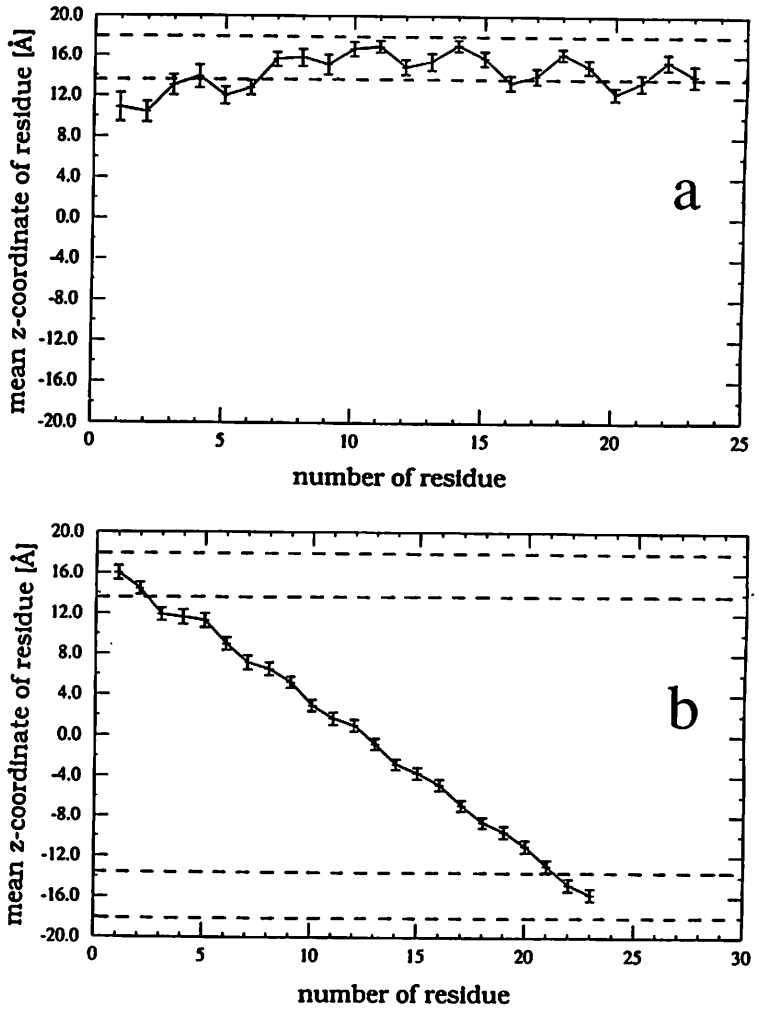

Figure 3. Mean z coordinates of residues for (a) magainin2 (a), (b) M2δ, and (c) $pf1$ and fd models.

simulation. Figure 3b displays the mean z coordinates of the M2δ residues. The result is very different from the magainin2 case. The residues of M2δ are uniformly distributed along the z coordinate, which lies across the membrane. The small mean deviation values indicate that the vertical motion of the residues is negligible after the transmembrane structure forms. Thus, M2δ forms a stable, transmembrane helix.

Taken together, these two examples demonstrate that on the basis of sequence information alone, this model can distinguish between transmembrane and surface adsorbed peptides.

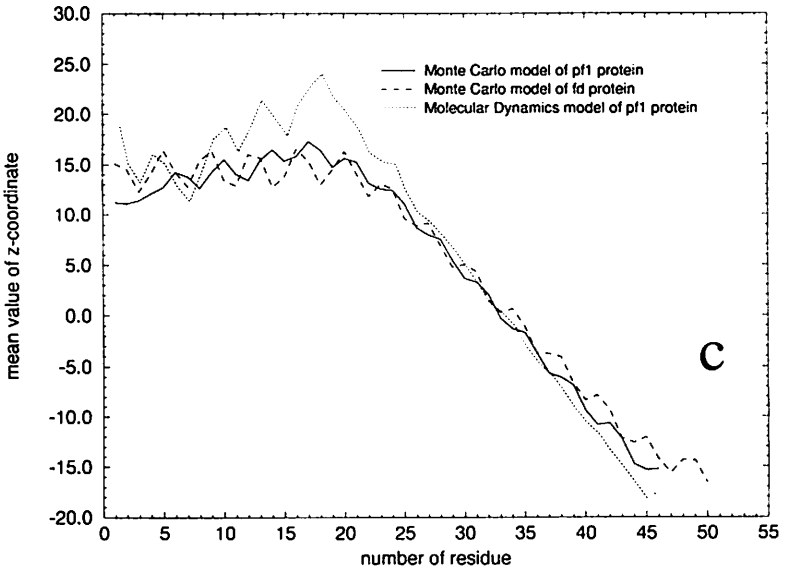

Figure 3. (Continued)

FILAMENTOUS BACTERIOPHAGES *Pf*1 AND *fd* COAT PROTEINS

Pf1 and fd represent two classes of filamentous bacteriophages identified on the basis of their X-ray diffraction patterns (Clark and Gray, 1989; Nakashima et al, 1975). Virions of these phages consist of a single stranded DNA ring coated by a protein layer. There are only two virion proteins, and the major protein (B) constitutes more than the 90% of the protein coat. All the major coat proteins for filamentous bacteriophages are relatively short (about 50 amino acids), with large percentages (close to 100%) of α-helical secondary structure. According to experimental data (Shon et al, 1991), in the membrane both proteins form structures containing transmembrane and interface adsorbed helices which are connected by a flexible linker. Thus, these sequences have been used in order to test the model on larger and more complicated protein structures. In our simulations, the behavior of both model pf1 and fd sequences are very similar; therefore, the process of insertion is presented here for the example of the fd coat protein. Figure 4 presents snapshots of the more important stages of the insertion process. Starting from a random conformation outside of the membrane, the peptide chain very rapidly adsorbs onto the surface of the membrane and forms a slightly distorted helical structure (Figure 4a). Then, the chain waits for a energy fluctuation that is sufficient to begin to transport the hydrophilic C-terminus of the protein across the membrane (Figure 4b). When the C-terminus crosses the membrane (Figure 4c), the final structure forms very quickly and remains stable during the remainder of the simulation (Figure 4d).

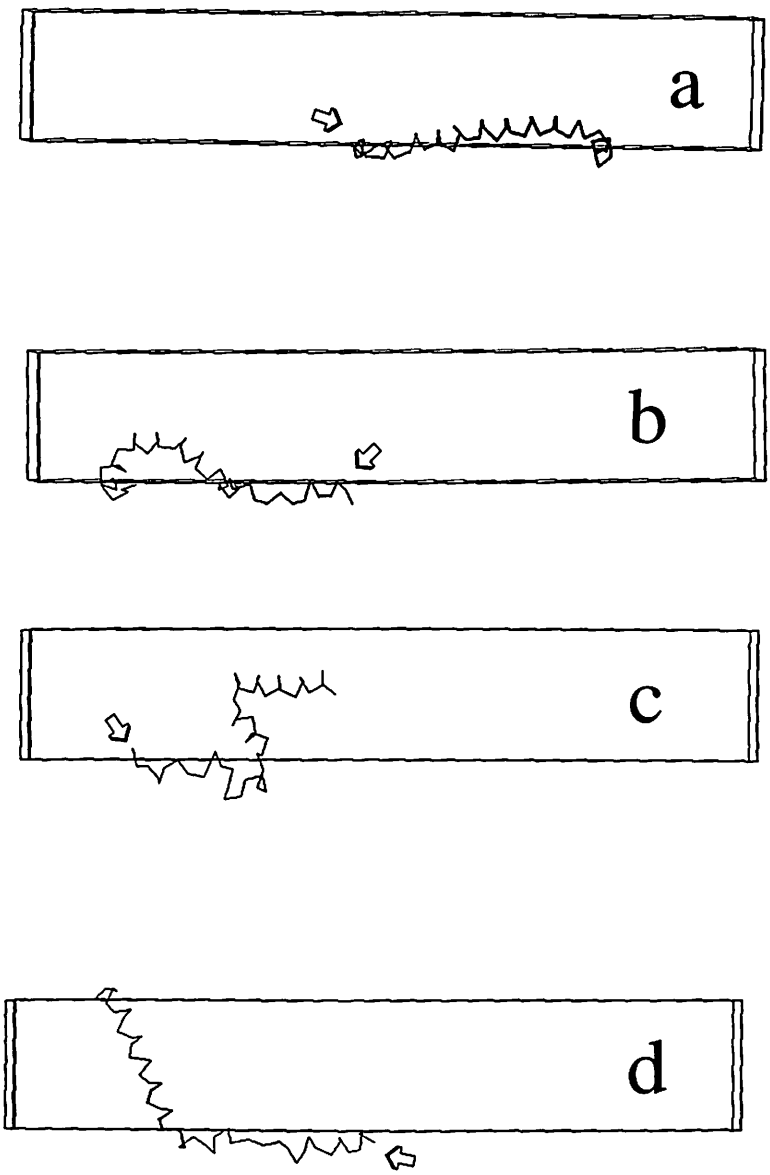

Figure 4. Snapshots from a typical simulation of the membrane insertion mechanism of the *fd* filamentous bacteriophage coat protein. The arrows point to the *N*-terminus of the chain.

In order to check the stability of the final state, a simulation ten times longer than usual ($5 \cdot 10^7$ MC steps) has been run. The simulation starts from the final, assembled conformation of the fd chain, with all parameters as in previous runs. In spite of the very fast lateral diffusion of the peptide in the membrane phase, both the orientation of the helices and the location of the

turn are stable during the simulation. This confirms that the final structure obtained during the simulation is in a deep free energy minimum.

The sensitivity of the model to temperature has been tested by simulating the peptide-membrane system at a temperature of 270K, instead of the usual 305K. The behavior of the system at the lower temperature is analogous to that at higher temperature. Insertion of the fd coat proteins into the membrane phase is not substantially more difficult at lower temperature; thus, over the examined temperature range, the final structure does not depend on temperature.

One of the important features of our model is the possibility of obtaining information about the position in the sequence of the transbilayer regions, adsorbed fragments, and turns. This information is not assumed in our model, because uniform helix propensities are employed for the entire chain. However, it could be possible that the predicted protein structure is an artifact of the geometry of the model membrane. To eliminate this possibility, a set of runs with different values of the membrane thickness have been undertaken. The results of simulations on systems in which the thickness of the lipid phase ranged from 21 Å to 27 Å show that the final structure of the model membrane protein does not change. Thus, we are confident that the position of the predicted turn and the orientation of the helical fragments does not depend on the membrane geometry. Rather, the structures of the membrane proteins predicted by the present model depend strongly on the hydrophobicity pattern of the protein sequence and do not depend on the simulation parameters (at least over the range we investigated).

Figure 3c shows the average values of the z coordinates of the C^{α} carbons of pf1 and fd coat proteins, obtained using the present method. In spite of their very low sequential homology, the similarity of the hydrophobic patterns implies a very similar structure. Both protein structures are predicted to have an interfacial adsorbed helix, which is 20 residues long, and a transbilayer helix stretching from residues 22 to residue 42 in pf1 and to residue 43 in fd. In Figure 3c, we also superimpose the mean z coordinates obtained by Roux and Woolf for pf1 (see chapter by Roux and Woolf in this volume). Their results, which are obtained using a different model and method, (all atom models sampled with molecular dynamics), are very similar to ours. The positions of the transmembrane helix and linker are very close in the Monte Carlo and molecular dynamics (MD) simulations. This confluence of results provides additional evidence that the results obtained in the present set of simulations provide some insight into the nature of the process of insertion of peptide chains into lipid membranes.

Translocation of Polymer Chains across a Curved Membrane

Another application of the Monte Carlo method to protein–membrane systems is due to Baumgärtner and Skolnick (1994a,b). Because they are designed to investigate the role of the membrane in the insertion process, these models use a more elaborate representation of the membrane phase at the expense of a very

simplified model of the protein chains, which lack any secondary structure. The cases of planar and curved membranes have been examined; the objective of this series of simulations is to determine whether membrane curvature exerts any effect on peptide translocation.

The Membrane Model

The membrane is comprised of an assembly of dumbbells tethered by one end to a penetrable interface. Because the dumbbells are tethered on both sides of the interface, they form a bilayer structure. The dynamics of the membrane is modeled by random moves of dumbbells in the monolayer (the length of each tether can vary from 0 to 0.7 model units) and by flip-flop moves, in which dumbbells can migrate from one layer to another. The last move is particularly important in the case of the model of a curved membrane, in which the planar densities of lipids are different on the cis, outside, and trans, inside, halves of the bilayer, and in which the density ratio is difficult to evaluate a priori. Every dumbbell is formed from two hard spheres of diameter $\sigma_L = 1.1$ (in the model units), connected by a rigid rod of length 1.32 (in the model units). The spheres represent the head and tail parts of lipids. In this realization, the bilayer is formed from 1000 dumbbells. The curved membrane is a sphere of radius 9.373. The equilibrium density of lipids in the cis half of the bilayer is 0.55, and in the trans half, it is 0.35. The difference in densities results in different average widths of the layers. The average width for the layers is 1.47 for the cis layer and 1.21 for the trans layer.

The Polymer Model

The polymer is modeled by a bead-spring chain, containing 20 hard spheres of diameter 0.69, connected by harmonic springs with a potential:

$$U(l) = K(l - l_0)^2 \tag{9}$$

with $K = 5$ and $l_0 = 0.7$. The average radius of such a chain is comparable with the width of one half of the model membrane.

The Potential

All the elements of the model system, lipid heads and tails and beads of the polymer chain, interact with a Morse potential of the form:

$$V(r) = \varepsilon[exp(-2\alpha(r - r_m)) - 2\exp(-\alpha(r - r_m))], \tag{10}$$

with parameters $\alpha = 4.0$ and $r_m = r_0 + \ln 2/\alpha$. The value of r_0 is set to be $(\sigma_L + \sigma_P)/2$ and $V(r_m) = -1.0$. The position of the polymer in the membrane phase is monitored by the mean penetration depth $\langle\Lambda\rangle$, defined as the average

distance of the center of mass of the polymer from the central plane of membrane. Positive values of $\langle \Lambda \rangle$ correspond to the cis side and negative values to the trans side of the sphere.

Results

The mean penetration depth value is calculated for flat and curved membranes, as a function of adhesion strength, γ/kT. At high temperatures, for $\gamma/kT < 0.45$, the polymer remains outside of both flat and curved membranes. For lower temperatures ($0.45 < \gamma/kT < 0.8$), the polymer in the curved membrane is spontaneously transported from the cis to the trans side of the membrane. This effect does not appear in the case of flat membrane, in which the polymer multiply crosses the bilayer. The total average interaction energy $\langle U/N \rangle$ between the polymer and the lipids, as a function of γ/kT, is very similar in the case of curved and flat membranes. Additionally, the average energy of a monomer as a function of its z coordinate is very similar in both cases. Therefore, the authors conclude that the process of transport is controlled by entropy and fluctuation effects. Indeed, if the average potential across the spherical bilayer is used in an effective medium model, the polymer spends roughly equal amounts of time on the trans and cis side of the curved bilayer. Furthermore, in the dumbbell model of the bilayer, the fluctuations of the orientational order parameter are two times larger on the trans than on the cis side. The polymer chain, after penetrating into the membrane, prefers to remain in the entropically more favorable, trans half of the bilayer. This portion of the membrane is of lower density; thus, the configurational entropy of the chain is larger. These simulations strongly indicate that there may be situations where explicit treatment of the lipid is extremely important, and thus, caution must be exercised in the use of effective medium membrane models. Certain physical properties (e.g., conformation) may be well described, but detailed insight into the transport process may require incorporation of the membrane molecules at some level of detail.

Model of Polymer Electrophoresis Across a Membrane

The model depicted above, with an additional electric field E, was used to simulate the process of electrophoresis of chain molecules across flat plane membranes. A short chain (length $N = 4$) of hard spheres, connected with harmonic bonds is used to model polymer chain. The electric field E is taken to be perpendicular to the flat membrane surface. The field effect on the dynamics of the polymer chain was simulated by modifying the dependence of the acceptance of a Monte Carlo move on the relative change of position during the move. The displacement from position r to r' is accepted, when $\exp[-E \cdot (r' - r)] > \eta$, where η was a pseudo-random number. The model of the membrane used in this simulation is analogous to the case of flat membrane depicted in the previous model and was formed by 1000 dumbbells, representing lipid molecules.

Results

As was the case in the previous simulations of Baumgärtner and Skolnick (1994b), the process of transportation was depicted by the position of center-of-mass of the polymer. The translocation time τ, the time needed for polymer to cross the membrane (from $\Lambda(0) > 0$ to $\Lambda(\tau) < 0$), was calculated as a function of field strength. According to authors, polymer electrophoresis across a fluctuating membrane can be related to a Kramer's problem, describing the average transition time for a Brownian particle escaping over a potential barrier. In the high field limit, the permeability, calculated as $\mu = \tau_0/\tau$ (where τ_0 is time of transition without the membrane present), is independent of the field. At low fields, the permeability exponentially decreases. In spite of its simplicity, the model gives some insights into the process of protein chain translocation across lipid membranes. Particularly interesting are experiments involving translocation of a chain across a frozen membrane; these emphasize the importance of fluctuation of the lipids in this process. Without lipid fluctuations, the polymer must randomly diffuse until it finds a hole through the entire membrane, a very unlikely process, whereas when the lipids fluctuate, the chain can burrow through the membrane.

Conclusions

There are a number of qualitative conclusions that emerge from the study of simplified models of peptide insertion into bilayers. First, for relatively small proteins, in accord with experimental evidence, peptide insertion in a bilayer can be spontaneous. Second, just as in globular proteins, the sequence determines the conformation of the model peptide. The location and orientation of a peptide with respect to a bilayer appears to be determined by the pattern of hydrophobic and hydrophilic residues. Thus, hydrophobic interactions play a very important part in the assembly process. Third, in addition to dictating the general features of the mechanism of spontaneous insertion, hydrogen bonds provide an additional, very important contribution to peptide stability in the bilayer. In the membrane interior, an isolated peptide can only hydrogen-bond to itself, whereas in the aqueous phase, a protein chain can readily exchange an internal hydrogen bond for one with solvent. Regardless of the absolute magnitude of a hydrogen bond, in the bilayer phase, the free energy cost of the absence of hydrogen bonds is very large. Thus, in agreement with the Engelman Steitz (1991) hypothesis of peptide insertion as modified by Jacobs and White (1989), in these simulations peptides first assemble helices on the membrane interface; then the almost fully assembled helices insert into the bilayer. Fourth, by building a model that incorporates the above features along with an experimentally determined transfer free energy scale, it is now possible to predict the orientation and conformation with respect to the bilayer of membrane associated peptides. It is especially encouraging that the results are in agreement with the more detailed models of Roux and Woolf in the present volume. This offers the prospect of simulating

more complicated systems that are beyond the range of computer resources for more detailed models, and/or the possibility of simulating a larger number of membrane peptide systems than would be possible for models at full atomic resolution. However, the treatment of peptides that are devoid of tertiary interactions is actually the simpler part of the more general problem of predicting the conformation of membrane proteins. What is still lacking is an effective free energy scale that describes intraprotein interactions in a lipid. Because of the paucity of experimentally determined crystal structures, pair potentials of mean force describing interactions of protein residues in a bilayer are unavailable. This makes it very difficult to predict whether a leucine would prefer to interact with another hydrophobic residue in the protein core or whether it would prefer to interact with lipid molecules instead. While we described a schematic lattice model that uses a conjectured pair potential to obtain some qualitative insights, this is still a far cry from an interaction scheme that is applicable to real proteins.

At this juncture, it is clear that the membrane can be treated as a structureless effective medium to model some properties of transmembrane proteins. However, the complete neglect of the structure of the bilayer can miss important physical effects, in particular, those associated with the entropy of the bilayer. For example, for a structureless polymer diffusing across a curved membrane, only when the lipid molecules are explicitly present does almost irreversible transport from the outside to the inside of the sphere occur. Granted the radii of curvature required to produce these effects are high, but they are not unphysical. Membranes having high curvature are found at various intracellular membrane bounded compartments (Alberts et al, 1994); thus, the effect of quasi-irreversible transport may have some physical basis. On the other hand, these simulations ignore the possibility of conformational transitions of the polymer as it diffuses through the bilayer. Clearly, what is required is the marriage of models having the possibility of secondary structure formation, as in the off-lattice membrane peptide model described above, with models that include a more detailed treatment of the membrane. Again, the level of detail required to treat the membrane is uncertain. Initially, to reduce the requisite simulation time, dumbbells might be used, but subsequently, more detailed models of membranes will have to be investigated. Overall, though, the use of reasonably simplified models of both the peptide and the membrane appear to be a very promising direction for providing not only qualitative insights into the peptide insertion process, but also for providing structural models of membrane peptides at moderate resolution.

REFERENCES

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994): *Molecular Biology of the Cell.* New York: Garland

Baumgärtner A, Skolnick J (1994a): Spontaneous translocation of a polymer across a curved membrane. *Phys Rev Lett* 74:2142–45

Baumgärtner A, Skolnick J (1994b): Polymer electrophoresis across a model membrane. *J Phys Chem* 98:10655–58

Bechinger B, Kim Y, Chirlian LE, Gessel J, Neumann JM, Montal M, Tomich J, Zasloff M, Opella SJ (1991): Orientations of amphiphilic helical peptides in membrane bilayers determined by solid-sate NMR spectroscopy. *J Biomol NMR* 1:167–73

Binder K (1984): *Applications of the Monte Carlo Method in Statistical Physics*. Heidelberg: Springer-Verlag

Clark BA, Gray DM (1989): A CD determination of the α-helix contents of the coat proteins of four filamentous bacteriophages: fd, IKe, Pf1 and Pf3. *Biopolymers* 28:1861–73

Engelman DM, Steitz TA (1991): The spontaneous insertion of proteins into and across membranes: The helical hairpin hypothesis. *Cell* 23:411–22

Gierash LM (1989): Signal sequences. *Biochemistry* 28:923–30

Godzik A, Kolinski A, Skolnick J (1993): Lattice representations of globular proteins: How good are they? *J Comp Chem* 14:1194–1202

Gregoret LM, Cohen FE (1990): Novel method for the rapid evaluation of packing in protein structures. *J Mol Biol* 211:959–74

Jacobs RE, White SH (1989): The nature of the hydrophobic binding of small peptides at the bilayer interface: implications for the insertion of transbilayer helices. *Biochemistry* 28:3421–37

Jacobs RE, White SH (1987): Lipid bilayer perturbations induced by simple hydrophobic peptides. *Biochemistry* 26:6127–34

Jahnig F (1983): Thermodynamics and kinetics of protein incorporation into membranes. *Proc Natl Acad Sci USA* 80:3691–95

Kolinski A, Skolnick J (1994): Monte Carlo simulations of protein folding. Lattice model and interaction scheme. *Proteins* 18:338–52

Lesser GJ, Lee RH, Zehfus MH, Rose GD (1987): Hydrophobic interactions in proteins. *Protein Engin* 14:175–9

Maduke M, Roise D (1993): Import of mitochondrial presequence into protein free phospholipid vesicles. *Science* 260:364–367

Milik M, Skolnick J (1993): Insertion of peptide chains into lipid membranes: An off-lattice Monte Carlo Dynamics model. *Proteins* 15:10–25

Milik M, Skolnick J (1992): Spontaneous insertion of peptide chains into membranes: A Monte Carlo model. *Proc Natl Acad Sci USA* 89:9391–95

Nakashima Y, Wiseman RL, Konigsberg W, Marvin DA (1975): Primary structure of side chain interactions of Pf1 filamentous bacterial virus coat protein. *Nature* 253:68–71

Rapoport TA (1992): Transport of proteins across the endoplasmic reticulum membrane. *Science* 28:931–36

Rapoport TA (1991): Protein transport across the endoplasmic reticulum membrane: Facts, models, mysteries. *FASEB J* 5:2792–98

Roseman MA (1988): Hydrophilicity of polar amino acid side chains is markedly reduced by flanking peptide bonds. *J Mol Biol* 200:513–22

Shon J, Kim Y, Colnago LA, Opella SJ (1991): NMR studies of the structure and dynamics of membrane-bound bacteriophage Pf1 coat protein. *Science* 252:1303–5

Singer SJ (1990): The structure and insertion of integral proteins in membranes. *Annu Rev Cell Biol* 6:247–96

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

Wolfe PB, Rice M, Wickner W (1985): Effects of two sec genes on protein assembly into the plasma membrane of *Escherichia coli*. *J Biol Chem* 260:1836–41