Deciphering the folding kinetics of transmembrane helical proteins

Enzo Orlandini*, Flavio Seno*†, Jayanth R. Banavar‡, Alessandro Laio§, and Amos Maritan§¶

*Istituto Nazionale Fisica della Materia (INFM), Dipartimento di Fisica, Università di Padova, Via Marzolo 8, 35131 Padova, Italy; ‡Department of Physics and Center for Materials Physics, The Pennsylvania State University, University Park, PA 16802; §INFM, International School for Advanced Studies, Via Beirut 4, 34014 Trieste Italy; and ¶The Abdus Salam International Center for Theoretical Physics (ICTP), Strada Costiera 11, 34100 Trieste, Italy

Nearly a quarter of genomic sequences and almost half of all receptors that are likely to be targets for drug design are integral membrane proteins. Understanding the detailed mechanisms of the folding of membrane proteins is a largely unsolved, key problem in structural biology. Here, we introduce a general model and use computer simulations to study the equilibrium properties and the folding kinetics of a C_α-based two-helix bundle fragment (comprised of 66 aa) of bacteriorhodopsin. Various intermediates are identified and their free energy are calculated together with the free energy barrier between them. In 40% of folding trajectories, the folding rate is considerably increased by the presence of nonobligatory intermediates acting as traps. In all cases, a substantial portion of the helices is rapidly formed. This initial stage is followed by a long period of consolidation of the helices accompanied by their correct packing within the membrane. Our results provide the framework for understanding the variety of folding pathways of helical transmembrane proteins.

Considerable effort has been expended to understand the dynamics of the folding and biological functionality of proteins. Whereas the behavior of small water-soluble globular proteins is reasonably well understood both experimentally and theoretically (1, 2), much less is known about membrane proteins (MP) (3–6) that cross biological membranes. Transmembrane proteins (TMP) are the most important and best studied class of MP (3, 4, 7). They are characterized by the presence in their primary structure of long segments (20–30 aa) of amino acids with a high degree of hydrophobicity. In the native structure, these correspond to the transmembrane segments that are inserted in the lipidic interior of the membrane (8). These segments are predominantly made up of α-helices and β-sheets. The stability of α-helices and β-sheets inside the membrane follow from the formation of hydrogen bonds between the backbone atoms—other possibilities are excluded within the apolar environment (3, 9).

Phenomenological models have proved to be powerful for interpreting experimental data. The most common of these is the *two-stage model* based on experimental evidence that the folding of TMP occurs in two stages. In the first stage, α-helices and β-sheets are formed with the full native state structure being formed in a distinct second stage (9). A more refined model (3) takes into account four main steps: partitioning, folding, insertion, and association. Recently, Pappu *et al.* (10) have used a potential smoothing algorithm to predict transmembrane helix packing in good accord with experimental data.

Milik and Skolnick (11, 12) have carried out careful Monte Carlo studies of the insertion of peptide chains into lipid membranes and have proposed a new hydropathy scale based on experimental data obtained by studying the interactions of tripeptides with phospholipid membranes (13) and the self-solvation effect in protein systems (14). Recently, Wimley and White (15) have designed transmembrane peptides that spontaneously insert across bilayers but yet have measurable mono-

meric water stability, opening the way for the determination of the thermodynamic cost of partitioning hydrogen-bonded peptide bonds into the membrane hydrocarbon core.

The Monte Carlo results of Milik and Skolnick (11, 12) are in good accord with Engelman and Steitz's (16) helical hairpin hypothesis further extended by Jacobs and White (17, 18). The unfolded chain first is adsorbed onto the membrane interface, driven mostly by the hydrophobic effect and electrostatic lipid–protein interactions (19–21). A polypeptide chain has a greater possibility, while anchored to the interface, of saturating its internal hydrogen bonds and forming helices. Such helical fragments have a greater propensity to subsequently diffuse into the lipid phase.

A detailed study of TMP has not yet been possible because little is known about the interactions between amino acids inside the membrane or between them and the lipid molecules. Here, we adopt a simple, yet powerful, strategy for attacking the folding properties of TMP that circumvents this shortcoming. Our approach is based on extensive studies of the folding of globular proteins that have underscored the important role played by the topology of the native state in controlling both the functionality and the main features of the folding process. Nature uses a rich repertory of 20 kinds of amino acids with sometimes major and at other times subtle differences in their interactions with the solvent and with each other in order to design sequences that fit the putative native state with minimal frustration (22). Thus, a fruitful and general strategy for the study of protein folding would be to extract information on the folding process directly from the topology of the native state.

Our study here focuses on the folding process by using a tractable approach (described in the *Appendix*) that bypasses the details of the complex interactions of the protein in the lipid environment by introducing effective potentials, induced by the presence of the membrane and the associated interface region, that stabilize the native state structure. The validity of the approach based on the native state topology, in the case of globular proteins, has been confirmed a posteriori from the agreement between theory and experimental findings (1, 23–31). The approach proposed here is similar in spirit and ought to be a tool and a guide for the difficult experimental situation of TMP (15). Our model allows a complete characterization of the thermodynamics and the dynamics of the full folding process. Because of the small number of degrees of freedom involved, the dynamics of the system can be simulated for the full folding process. Moreover, the free energies of the most relevant intermediate states and free energy profiles along the reaction paths connecting them can be explicitly calculated by thermo-
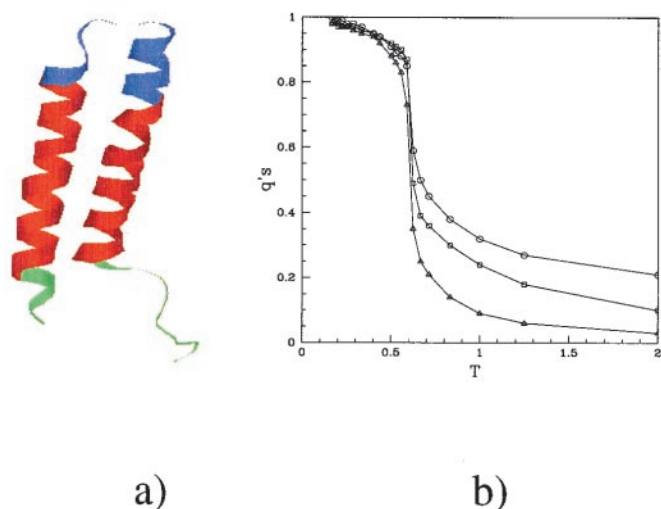
BIOPHYSICS

Fig. 1. Structure and thermodynamics of the helical transmembrane protein. (*a*) Ribbon representation of the two-helix fragment of bacteriorhodopsin formed by the first 66 aa. The part inside the membrane (determined by using the neural network learning algorithm available at http://www.embl-heidelberg.de/Services/sander/predictprotein/) is shown in red, the part above (below) the membrane in blue (green). (*b*) Average equilibrium fraction of native contacts outside, q_b (○), inside, q_m (□), and across, q_s (△), the membrane as a function of the temperature T. All these quantities are expressed in energy unit of ε (see *Appendix*). The folding transition temperature T_C when all the curves cross the value 1/2 is approximately 0.6. This value is in accord with the temperature of the heat capacity maxima.

dynamic integration (see *Appendix*). Thus, the model is able to quantitatively discriminate between the possible reaction paths envisaged for the insertion process of TMP across the membrane (3), a feature that is not an obvious consequence of the structure of the model. Where there is overlap, our model captures the qualitative features of the earlier simulations of Milik and Skolnick (11, 12).

The TMP we considered is made up of the first 66 aa of bacteriorhodopsin consisting of two α-helices (Fig. 1*a*). It has been shown that the first two helices of bacteriorhodopsin can be considered as independent folding domains (32). Furthermore, the side-by-side interactions between transmembrane helices play a key role in the stabilization of the protein structure (33).

Our studies were carried out using a Monte Carlo algorithm that has been proven to be extremely efficient for interacting heteropolymers (see *Appendix*). The behavior of the structural similarity between the system equilibrated at temperature T (measured in dimensionless units) and the native state is shown in Fig. 1*b* in terms of the average fraction of native state contacts as a function of T and partitioned depending on their positions with respect to the membrane. The three curves correspond respectively to the average fraction of native contacts inside (q_m), outside (q_b) and across (q_s) the membrane (see *Appendix*). All these curves, well separated at high T, collapse for T below the transition temperature $T_C \approx 0.6$, indicating a cooperative effect in the folding. On monitoring the free energy as a function of the energy around T_C, one observes additional local minima (besides those corresponding to the unfolded and folded states), suggesting the presence of an intermediate.

The intermediate is characterized by having the two helices almost completely formed but not yet correctly inserted across the membrane. A metastable state in which the protein exists at the membrane interface ought to be expected on general

grounds. Indeed a generic heteropolymer with hydrophobic and hydrophilic aminoacids, of which a TMP is a particular case, has a favorable conformation that is localized near surfaces between two selective media (the outside and the inside part of the membrane in the present case) (34, 35). At not-too-high temperatures, the gain in energy to place hydrophobic/hydrophilic protein segments in their preferred environment compensates the entropy loss for being localized at the interface with respect to remain in the bulk phase. Thus, even though our model does not explicitly contain information on the character of the amino acids, it is able to predict this feature.

The presence of these extra minima suggests that nonconstitutive membrane proteins would fold with multistate kinetics corresponding to on-pathway intermediates. To establish their nature of and their influence on the dominant folding pathways, we have performed a detailed analysis of the folding kinetics. Each independent kinetic folding simulation was started with the equilibrated denatured state at $T^* = 2.5$. The protein is placed initially outside the membrane in the interface region (3) at a distance comparable to the average size of the denatured protein and then suddenly quenched to a temperature ($T = 0.4$) well below the transition temperature. This case simulates the folding kinetics of nonconstitutive membrane proteins, i.e., proteins that do not need a translocon providing a "tunnel" through which the protein is injected into the lipid bilayer. Folding to the native state occurs mainly through the states depicted in Fig. 2*a* with the dominant pathways shown in Fig. 2*b*.

In all the pathways, the system goes from the unfolded state U to state HI in which 80% of the secondary structure is formed (see q in Fig. 3*c*) and disposed horizontally along the interface. The free energy of this state (measured with respect to the free energy of the fully folded state) is ≈ 2.4 T_C. This state corresponds to the formation of approximately 70% of the membrane contacts. The average time τ_{HI} to reach state HI is of the order of 500 Monte Carlo steps (see Figs. 3 and 4; each Monte Carlo step corresponds to 50,000 attempted local deformations). State HI turns out to be an obligatory on-pathway intermediate of the folding kinetics for nonconstitutive MP in agreement with the general argument mentioned above. Once the protein reaches state HI, it undergoes a relatively slow process of self-arrangement in order to insert and assemble the secondary structures across the membrane. This process is the rate-limiting step of the folding process, because it involves the translocation, through the lipidic layer, of a substantial number of hydrophilic residues. Among the possible pathways, starting from HI, the most frequent (60% of the cases) and the fastest turn out to be $U \rightarrow HI \rightarrow HV \rightarrow N$. A quantitative characterization of this dominant pathway is presented in Figs. 3 (for a single folding process) and 4 (as an average over 40 folding processes). The intermediate HV is characterized by having one α helix inserted across the membrane and is reached in an average period corresponding to a significant fraction of the total folding time (see Fig. 3). The free energy in this state is ≈ 0.98 T_C. The free energy barrier between HI and HV is at ≈ 4.31 T_C (hence, the rate constant of the transition $HI \rightarrow HV$ is proportional to $k_{HI \rightarrow HV} = \exp(-(4.31 - 2.4)T_C/T)$). The full free energy profile versus a reaction coordinate is shown in Fig. 5. The last part of the folding process corresponds to the insertion of the second helix and the assembly of the two secondary structures into the native state structure. This process lasts approximately one-third of the folding time along the pathway $U \rightarrow HI \rightarrow HV \rightarrow N$. The quasistatic free energy barrier between HV and the folded state is ≈ 1.66 T_C. The rate constant of the transition $HV \rightarrow N$ is therefore proportional to $\exp(-(1.66 - 0.98)T_C/T)$. These results are consistent with the time scales observed in the unconstrained folding dynamics. At the end, the protein is completely packed [q_m saturates to 1 (Figs. 3*a* and 4*a*)] and the helices are correctly positioned across the
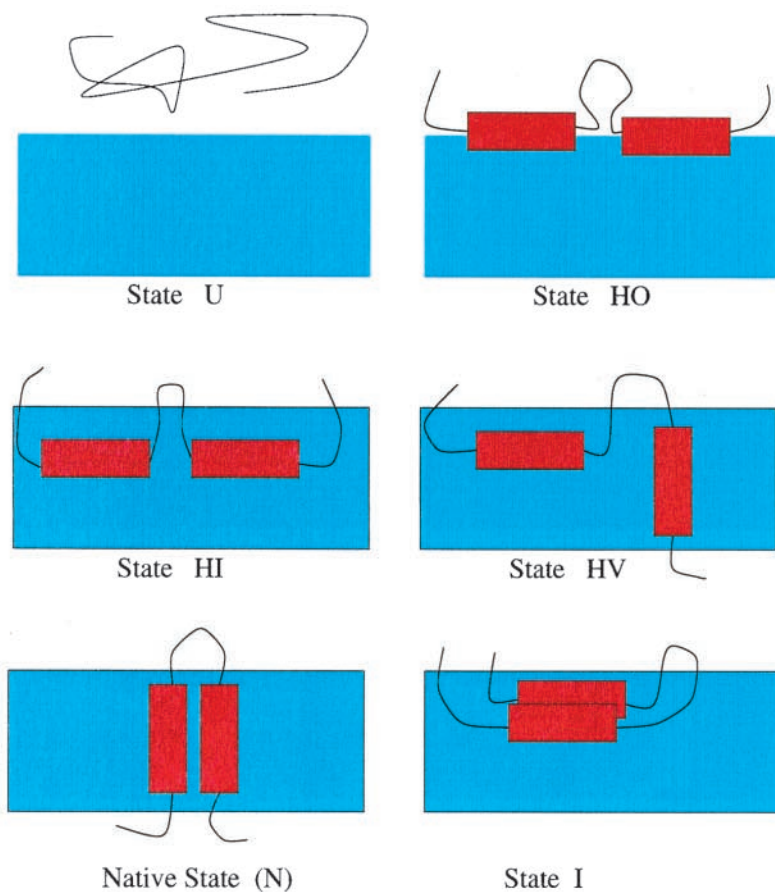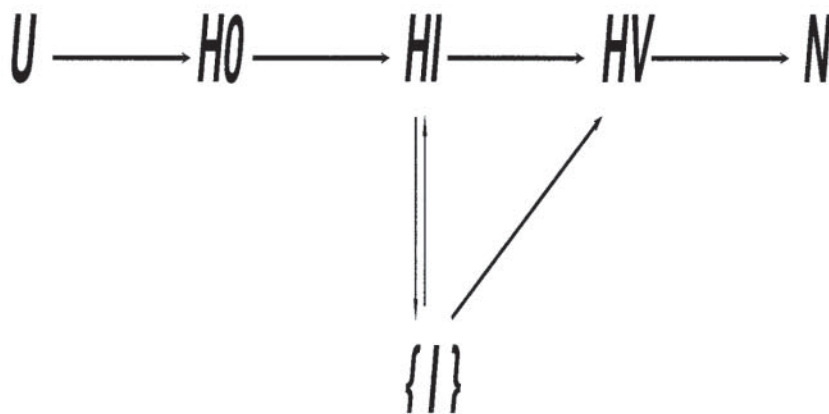
Fig. 2. Schematic representation of states encountered by nonconstitutive proteins during the folding process. (*a*) The red cylinders denote α-helices that reside within the membrane in the native state. The region inside the membrane is in turquoise, whereas the rest represents the interface region (3) in which the folding process starts. State *U* denotes the denatured state of the protein, *HO* is a state in which the helices have been formed but are not yet inside the membrane, whereas *HI* corresponds to a similar state but with the helices completely embedded in the membrane without any interhelical contacts. Usually the helices form and enter into the membrane separately. *HV* denotes an obligatory intermediate and *N* depicts the native state. The state {*I*} represents an ensemble of long-lived conformations in which helices are formed inside the membrane with several interhelical contacts, but with the two α-helices still incorrectly positioned. This conformation differs in terms of packing efficiency of the helices. The state {*I*} is not obligatory for the folding kinetics. (*b*) The schematic pathways to the native state are shown. In the most directed path, the entropy decreases on going from *U* to *N*. From *HI* to *HV* the entropy loss of one helix is not compensated by a corresponding energy gain until both helices become vertical. This is the principal origin of the high free energy barrier between the state *HI* and the native state.
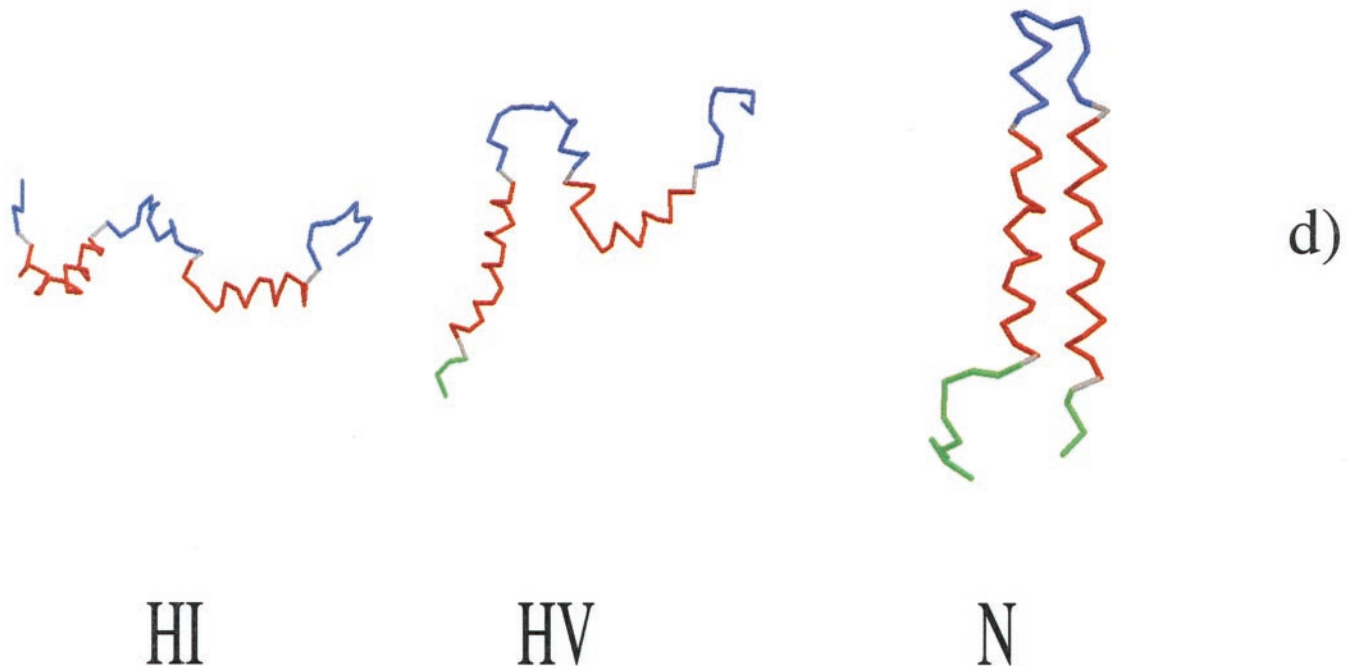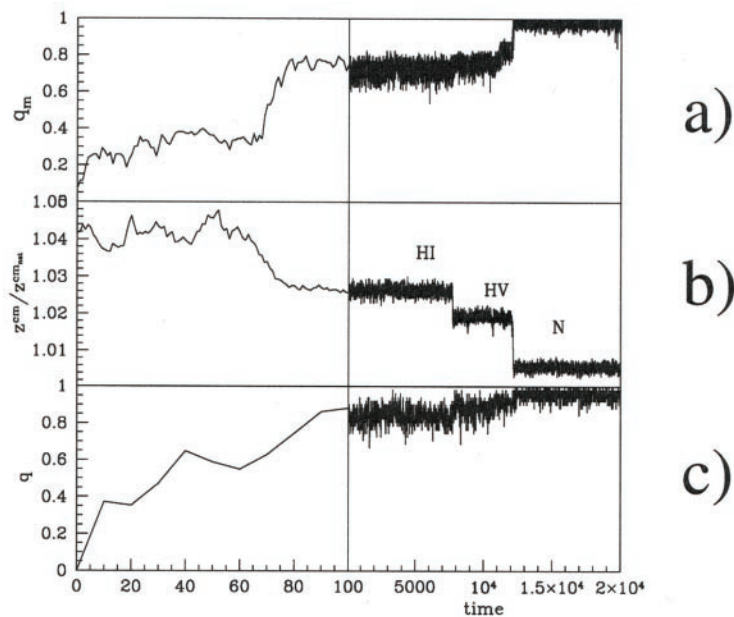
Fig. 3. Typical time dependence of different parameters as a function of the Monte Carlo steps for the pathway $U \to HI \to HV \to N$. Fraction of native contacts inside the membrane (*a*), normalized *z* coordinate of the center of mass of the protein (with respect to that of the native state conformation) (*b*) and overall fraction of native helical contacts (*c*). Each Monte Carlo step corresponds to 50,000 attempted local deformations. The transition from state *HI* to state *HV* is signaled by a sharp jump of the position of the center of mass. Note that there is no perceptible sign of this transition in terms of newly formed native contacts. Most of the helical contacts are formed in the early stages of the folding. This fraction does not significantly increase until helices correctly assemble and the interhelical contacts are formed. The $HV \to N$ transition is reached by a progressive zippering of the horizontal and vertical helices. This zippering is usually very quick (few Monte Carlo steps) and is only slightly slowed down (see the plateau corresponding to $q_m \approx 0.9$ in *a*) when the trajectory passes through somewhat deformed conformations. (*d*) Protein conformations at different times during the folding. The colors red, green, and blue have the same significance as in Fig. 1*a* with the grey bonds being ones crossing the membrane.

membrane (note the second jump in the *z* coordinate of the center of mass in Figs. 3*b* and 4*b*).

Much slower dynamics can occur when nonobligatory intermediates are visited by the system. These long-lived states ({*I*} in Fig. 2*a*) involve a distribution of misfolded regions that trap the system and are characterized by having most of the inter-

helical contacts formed (assembly of the secondary structures) but with the two α-helices still incorrectly positioned. Note, for example, that in states {*I*}, only transmembrane contacts and some contacts outside the membrane are misplaced and they account for only a small fraction of the native state energy. For this reason, in the states {*I*}, the free energy is $\approx 1.44\, T_C$, only
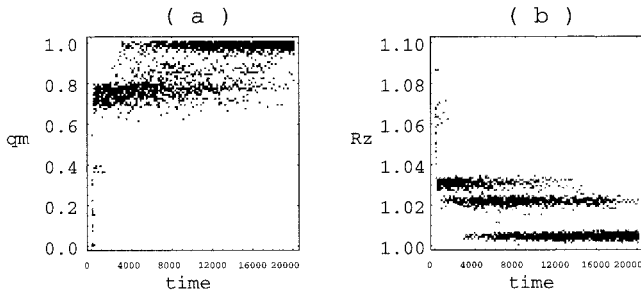
Fig. 4. Distribution of the fraction of native contacts inside the membrane (*a*) and of the normalized *z* coordinate of the center of mass ($Rz = z^{cm}/z^{cm}_{nat}$) (*b*). The data were obtained by using 40 independent kinetic simulations with pathway $U \rightarrow HI \rightarrow HV \rightarrow N$. The grey-scale distribution indicates the probabilities at given times: darker points denote higher probability.

slightly higher than the free energy of *HV*. The folding can proceed from {*I*} either by disentangling the two helices and passing through the obligatory intermediate *HV*, or by the simultaneous translocation through the membrane of the two helices. These processes, however, entail the crossing of a big free energy barrier ($\approx 5.18\ T_C$ for the first process and 6.1 for the second) and happen with low probability. Indeed, at sufficiently low temperatures, the loss in energy of the interhelical contacts is not compensated by the gain in the configurational entropy because of the uncoupling of the α-helices. Thus, below the folding temperature, *I*-states act as trapping regions for the system and when trapped, the protein spends most of the time during folding in this state.

In summary, we have presented detailed calculations of helical transmembrane proteins leading to a vivid picture of the folding process. Our strategy relies on the dominant role played by the topology of the native state structure and by the effective geometry imposed by the membrane and provides a picture that would be expected to be quite accurate for well designed sequences that are a good fit to the target native state conformation. It is interesting to note that, with our choice of the parameters, the pathway in which the helices assemble outside the membrane and are inserted later is unlikely to occur.

Models based on the topology of the native state structure have been remarkably successful (25–28) in correctly describing the main features of the folding process determined in experiments (1, 23, 24, 29–31) for various globular proteins. A similar approach has been generalized here to the almost virgin field of transmembrane proteins where experiments are rather difficult (3, 6, 15). Our findings do not depend on the precise values of the ε parameters introduced in the model underscoring the robustness of the results. Our approach predicts a folding process involving multiple pathways with a dominant folding channel. The simplicity of our model allows for a quantitative description of all the pathways because we can monitor the correct/incorrect formation of native contacts and compute free energy profiles. Further details not captured by the present approach arising from amino acid-specific interactions among themselves, with the solvent, and in particular with the interior of the membrane may of course change the quantitative nature of the results. However, our model, which captures the bare essentials of a membrane protein, ought to provide a zeroth-order picture of the folding process. Also, as experimental data becomes available, the results could be benchmarked with models of this type to glean the other factors that matter.

Appendix

We represent the residues of the membrane protein as single beads centered in their C_α positions. Adjacent beads are tethered together into a polymer chain by a harmonic potential with the average $C_\alpha - C_\alpha$ distance along the chain equal to 3.8 Å. The membrane is described simply by a slab of width $w = z_{max} - z_{min} = 26\ \mathring{A}$. Two nonbonded residues (i, j) form a contact if their distance is less then 6.5 Å. In the study of globular proteins, the topology of the native state is encoded in the contact map giving the pairs (i, j) of nonbonded residues that are in contact. Here, in addition, the locations of such pairs with respect to the membrane becomes crucial. The contacts are divided into three classes: *membrane contacts* where both i and j residues are inside the membrane, *interface contacts* with i and j in the interface region (3) outside the membrane, and *surface contacts* with one residue inside the membrane and the other outside. Thus, a given protein conformation can have a native contact but improperly placed with respect to the membrane (*misplaced native contact*).
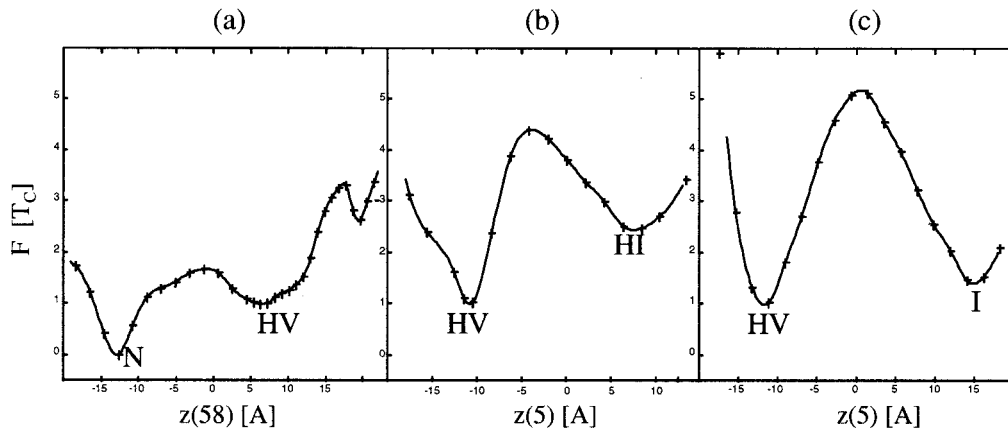
Fig. 5. Free energy profiles along three reaction coordinates at $T = 0.85\ T_C$. The continuous lines are spline fits to the free energy data (crosses). To obtain free energy differences between two states, we estimate the reversible work that has be done to go from one state to the other. For this purpose, we fix the *z* coordinate of a specific residue in order to compute the canonical average of the force and then apply Eq. **2** (see *Appendix*). The free energy of the native state is defined to be equal to 0. (*a*) Free energy as a function of the *z* coordinate of the 58th residue ($z = 0$ corresponds to the middle of the membrane) starting from *HV*; this forces the second helix to cross the membrane as the protein goes from *HV* to *N*; the local minimum at $z \approx 20$ corresponds to a state topologically equivalent to *HV*, with the helix containing the 58th residue fully formed on the membrane interface but without any contact with the first helix (in *HV* some of the interhelices contacts are already formed); (*b*) the 5th residue is translocated across the membrane with the protein starting from state *HI* and proceeding to *HV*; (*c*) the same as in *b*, but the initial state is *I* (see Fig. 2*a*).

The crucial interaction potential between nonbonded residues (i, j) is taken to be a modified Lennard–Jones 12–10 potential:

$$\Gamma(i, j)\left[5\left(\frac{d_{ij}}{r_{ij}}\right)^{12} - 6\left(\frac{d_{ij}}{r_{ij}}\right)^{10} \right] + 5\Gamma_1(i, j)\left(\frac{d_{ij}}{r_{ij}}\right)^{12}. \quad [1]$$

The matrices $\Gamma(i, j)$ and $\Gamma_1(i, j)$ encode the topology of the TMP in the following way: if (i, j) is not a contact in the native state $\Gamma(i, j) = 0$, $\Gamma_1(i, j) = 1$; if (i, j) is a contact in the native state but not at the proper location (i.e., a misplaced contact) $\Gamma(i, j) = \varepsilon_1$, $\Gamma_1(i, j) = 0$; if (i, j) is a native state contact in the proper region $\Gamma(i, j) = \varepsilon$, $\Gamma_1(i, j) = 0$. This model is intended to describe the folding process in the interface and in the membrane region. Our interaction potential {similar in spirit to a well known model (36) for globular proteins [see also other approaches that model helix formation (37, 38)]} assigns two values to the energy associated with the formation of a native contact, ε and ε_1.

The model captures the tendency to form native contacts. In addition, in order to account for the effective interactions between the membrane and the protein, the model assigns a lower energy, $-\varepsilon$, to the contact that occurs in the same region as in the native state structure compared to $-\varepsilon_1$ when the contact is formed but in the wrong region of space. This feature proves to be crucial in determining the mechanism of insertion of the protein across the membrane in order to place all native contacts in the same regions as in the native state. Even though the interaction potential is simple and intuitively appealing, it is not possible to simply guess (without detailed calculations) the folding mechanism and quantitatively determine the probability of occurrence of the various folding pathways (3).

When $\varepsilon = \varepsilon_1$, the protein does not recognize the presence of the interface-membrane region and the full rotational symmetry is restored (the system behaves like a globular protein). The difference in the parameters $(\varepsilon - \varepsilon_1)$ controls the amount of tertiary structure formation outside the membrane. When the difference is small, the protein assembles almost completely outside the membrane and the insertion process would be diffusion-limited. Our results are independent of the precise values of the energy parameters ε and ε_1 ($\varepsilon > \varepsilon_1$) as long as they are not too close to each other.

We report here the results of simulations with $\varepsilon_1 = 0.1$ and $\varepsilon = 1$. r_{ij} and d_{ij} are the distance between the two residues (i, j) and their distance in the native configuration, respectively. In order to account for the chirality of the TMP, a potential for the pseudodihedral angle α_i between the C_α atoms in a helix corresponding to four successive locations is added, which biases the helices to be in their native state.

The thermodynamics and the kinetics of the model were studied by a Monte Carlo method for polymer chains allowing for local deformations. The efficiency of the program (usually low for continuum calculations) has been increased by full use of the link cell technique (39) and by the multiple Markov chain method, a new sampling scheme, which has been proven to be particularly efficient in exploring the low-temperature phase diagram for polymers (40). In our simulation, 20 different temperatures ranging from $T = 2$ to $T = 0.17$ have been studied. The free energy is calculated by reweighting the different temperatures with the Ferrenberg–Swendsen (41) algorithm.

The free energy difference $\mathcal{F}_B - \mathcal{F}_A$ between two states A and B has been estimated as the reversible work that has to be done in order to go from A to B. Hence, denoting by $\mathbf{x}(\lambda)$ a reaction coordinate connecting A and B (for $\lambda = 0$ and $\lambda = 1$ the system is in A and B, respectively), and by $\langle \cdot \rangle_\lambda = \langle \delta(\mathbf{x} - \mathbf{x}(\lambda)) \cdot \rangle$, the canonical average at fixed reaction coordinate

$$\mathcal{F}_B - \mathcal{F}_A = \int_0^1 d\lambda \langle \mathbf{F} \rangle_\lambda . \frac{d\mathbf{x}(\lambda)}{d\lambda}$$

$$\simeq \sum_i \langle \mathbf{F} \rangle_\lambda . \frac{d\mathbf{x}(\lambda)}{d\lambda}\bigg|_{\lambda = \frac{\lambda_i + \lambda_{i+1}}{2}} (\lambda_{i+1} - \lambda_i), \quad [2]$$

where $\mathbf{F}$ is the force and $\{\lambda_i, i = 1, \ldots\}$ is a suitably dense partition of the interval $(0, 1)$. The average value $\langle \mathbf{F} \rangle_{\lambda_i}$ at each λ_i is computed by a long (more than 5,000 steps) Monte Carlo run performed with dynamics satisfying the constraint $\mathbf{x} = \mathbf{x}(\lambda_i)$. The free energy differences obtained with this method are accurate to within ≈ 0.1 T_C for the various states, whereas the free energy barriers are accurate within ≈ 0.5 T_C. This error takes into account possible hysteresis effects attributable to the finite simulation time.

1. Fersht, A. R. (1999) *Structure and Mechanism in Protein Science* (Freeman, New York).
2. Karplus, M. & Sali, A. (1995) *Curr. Opin. Struct. Biol.* **5,** 58–73.
3. White, S. H. & Wimley, W. C. (1999) *Ann. Rev. Biophys. Biomol. Struct.* **28,** 319–365.
4. Ostermeier, C. & Michel, H. (1997) *Curr. Opin. Struct. Biol.* **7,** 697–701.
5. von Heijne, G. (1996) *Prog. Biophys. Molec. Biol.* **66,** 113–139.
6. Booth, P. J. (1997) *Folding Design* **2,** R85–R92.
7. Biggin, P. C. & Sansom, M. S. P. (1999) *Biophys. Chem.* **76,** 161–183.
8. Deber, C. M. & Goto, N. K. (1996) *Nat. Struct. Biol.* **3,** 815–818.
9. Popot, J. L. & Engelman, D. M. (1990) *Biochemistry* **29,** 4031–4036.
10. Pappu, R. V., Marshall, G. R. & Ponder, J. W. (1999) *Nat. Struct. Biol.* **6,** 50–55.
11. Milik, M. & Skolnick, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9391–9395.
12. Milik, M. & Skolnick, J. (1993) *Proteins Funct. Struct. Genet.* **15,** 10–25.
13. Jacobs, R. E. & White, S. H. (1987) *Biochemistry* **26,** 6127–6134.
14. Roseman, M. A. (1988) *J. Mol. Biol.* **200,** 513–522.
15. Wimley, S. C. & White, S. H. (2000) *Biochemistry* **39,** 4432–4442.
16. Engelman, D. M. & Steitz, T. A. (1981) *Cell* **23,** 411–422.
17. Jacobs, R. E. & White, S. H. (1989) *Biochemistry* **28,** 3421–3427.
18. Jacobs, R. E. & White, S. H. (1986) *Biochemistry* **25,** 2605–2612.
19. Pinheiro, T. J. T., Elöve, G. A., Watts, A. & Roder, H. (1997) *Biochemistry* **36,** 13122–13132.
20. Rankin, S. E., Watts, A. & Pinheiro, T. J. T. (1998) *Biochemistry* **37,** 12588–12595.
21. Bryson, E. A., Rankin, S. E., Carey, M., Watts, A. & Pinheiro, T. J. T. (1999) *Biochemistry* **38,** 9758–9767.
22. Bryngelson, J. D. & Wolynes, P. G. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7524–7528.
23. Fersht, A. R. (1995) *Curr. Opin. Struct. Biol.* **5,** 79–84.
24. Fersht, A. R. (1997) *Curr. Opin. Struct. Biol.* **7,** 3–9.
25. Micheletti, C., Banavar, J. R., Maritan, A. & Seno, F. (1999) *Phys. Rev. Lett.* **82,** 3372–3375.
26. Munoz, V. & Eaton, W. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11311–11316.
27. Galzitskaya, O. V. & Finkelstein, A. V. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11299–11304.
28. Alm, E. & Baker, D. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11305–11310.
29. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M. & Dobson, C. M. (1999) *Nat. Struct. Biol.* **6,** 1005–1009.
30. Martinez, J. C. & Serrano, L. (1999) *Nat. Struct. Biol.* **6,** 1010–1016.
31. Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I. & Baker, D. (1999) *Nat. Struct. Biol.* **6,** 1016–1024.
32. Kahn, T. W., Sturtevant, J. M. & Engelman, D. M. (1992) *Biochemistry* **31,** 8829–8839.
33. Kahn, T. W. & Engelman, D. M. (1992) *Biochemistry* **31,** 6144–6151.
34. Garel, T., Huse, D. A., Leibler, L. & Orland, H. (1989) *Europhys. Lett.* **8,** 9–12.
35. Maritan, A., Riva, M. P. & Trovato, A. (1999) *J. Phys. A Math. Gen.* **32,** L275–L280.
36. Taketomi, H., Ueda, Y. & Go, N. (1975) *Int. J. Pept. Protein Res.* **7,** 445–459.
37. Guo, Z. & Thirumalai, D. (1996) *J. Mol. Biol.* **263,** 323–343.
38. Takada, S., Luthey-Schulten, Z. & Wolynes, P. G. (1999) *J. Chem. Phys.* **110,** 11616–11629.
39. Geroff, I., Milchev, A., Binder, K. & Paul, W. (1993) *J. Chem. Phys.* **98,** 6256–6539.
40. Tesi, M. C., van Rensburg, E. J., Orlandini, E. & Whittington, S. G. (1996) *J. Stat. Phys.* **29,** 2451–2463.
41. Ferrenberg, A. M. & Swendsen, R. H. (1989) *Phys. Rev. Lett.* **63,** 1195–1198.