Peptides at Membrane Surfaces and their Role in Prebiotic Evolution

Andrew Pohorille[1,2], Michael A. Wilson[1,2] and Christophe Chipot[3]

[1]Biomolecular and Cellular Modeling Program
NASA Ames Research Center, Moffett Field, CA 94035

[2]Department of Pharmaceutical Chemistry,
University of California, San Francisco, CA 94143

[3]Laboratoire de Chimie Théorique
Unité Mixte de Recherche CNRS 7565
Université Henri Poincaré – Nancy I, B.P. 239
54506 Vandœuvre–lès–Nancy, France

July 12, 2002

Abstract

Protocells had to transport ions and organic matter across membranes separating the interior of the cell from the environment, capture and utilize energy and transduce environmental signals. In a series of detailed, molecular-level computer simulations we show how these peptides in contact with membranes can acquire ordered structures and functions. We have investigated the stability of a simple α-helical peptide containing Leucine (L) and Serine (S) of the form $(LSLLLSL)_3$ in a model membrane system. The parallel in-plane state is the most stable configuration. The transmembrane state is metastable, and about 15 kcal mol^{-1} is required to insert the peptide into the membrane. We investigated dimers of both $(LSLLLSL)_3$ and glycophorin A, and show how the free energy of helix association can, at least partially, offset the free energy of insertion. We have also investigated the transmembrane pore of the influenza M2 protein. This aggregate of four

identical alpha-helices, each built of 25 amino acids, forms an efficient and selective voltage-gated proton channel. Our simulations explain the gating mechanism, which can involve strands of hydrogen-bonded water through the pore or proton transfer through tautomerization of protein residues. The channel can be re-engineered to act as a simple proton pump.

Key Words: peptides, peptide folding at interfaces, peptide-membrane interactions, peptide-petide association, Influenza M2 proton channel, protocellular functions.

Introduction

The emergence of vesicles — closed, spheroid assemblies of amphiphilic molecules enclosing an aqueous medium — was a likely early step in the evolutionary pathway from inanimate matter to the simplest living cells. Under proper conditions, vesicles form spontaneously from an aqueous solution of amphiphiles. No energy input or catalyst is required. The possible sources of amphiphiles on the primitive earth could have been terrestrial or extraterrestrial; vesicles have been shown to form from simple organic compounds, such as carboxylic acids, [5] the material extracted from the Murchison meteorite [34] and simulated cometary mantles. [37] Other proposed sites of protobiological evolution, such as clays or pyrites, are seemingly unrelated to modern cellular structures.

Vesicles became the precursors to true cells, or protocells, by acquiring the capabilities needed to survive and reproduce. Even the simplest protocells must have possessed such ubiquitous capabilities as transport of ions and nutrients across cell walls, capture and utilization of energy, and synthesis of molecules necessary for self-maintenance and growth. In modern organisms, these functions are carried out by proteins. Probably the most parsimonious assumption that can be made is that peptides were the protobiological precursors to proteins. Peptides may have emerged prior to, or independently of, RNA molecules, or might have been synthesized from a primitive RNA genome. In either scenario, their existence was an essential step leading to the emergence of contemporary cellular metabolism.

The assumption about the protocellular role of peptides is supported by the findings that amino acids could have been synthesized on the early

earth [21, 65] or delivered from extraterrestrial sources. [30] In particular, a model for the synthesis and polymerization of amino acids from simple precursors was recently proposed. [96, 97] Its essential features are that all reactions occur in the same environmental conditions (one-pot chemistry), the underlying chemistry resembles modern metabolism, and a mechanism for autocatalytic feedback is provided. In another model, membrane-assisted polycondensation of amino acids and peptides was demonstrated. [61] Small peptides have also been shown to form auto-catalytic, self-correcting networks. [56, 100] Under appropriate conditions, these networks can select for peptides built of amino acids of the same chirality. [81] Furthermore, small, functional proteins have been selected from random-sequence libraries with frequencies that appear to be similar to those observed for equivalent RNA libraries. [90]

On the basis of their location in the cell, proteins can be divided into two broad types — water soluble proteins, which reside in the cytoplasm, and membrane proteins. This division is useful because these two types of proteins have different structural properties, functions and evolutionary histories. Integral membrane proteins perform such essential cellular functions as transport of ions, nutrients and waste products across cell walls, transduction of environmental signals, regulation of cell fusion, recognition of other cells, energy capture and its conversion into high-energy compounds. In fact, 30-40% of genes in modern organisms codes for membrane proteins. [18, 94] In many instances, no alternative means of performing membrane-related functions are known, or even have been postulated.

Although contemporary membrane proteins can be quite complex, their transmembrane fragments are usually remarkably simple. The most common structural motif for these fragments is a bundle of α-helices, [7, 66] but occasionally is a β-barrel. [89] This is in contrast to water-soluble proteins, which exhibit a much wider range of structural motifs. Moreover, some membrane proteins retain their functionality even if a large fraction of the protein is removed, [69] whereas water-soluble proteins almost never do so. Another protobiologically relevant property is that some simple, α-helical synthetic or natural membrane peptides aggregate spontaneously and form functional complexes exhibiting sequence-dependent specificity. [2, 7, 8] Finally, short membrane peptides can acquire well-defined secondary structures, such as α-helices, β-strands or β-turns, [7, 35, 50, 77] whereas water-soluble peptides of similar length are typically disordered. These properties of membrane peptides suggest that their associations, especially in α-helical conformations, are

excellent candidates for mediators of transmembrane functions in protocells.

The conceptual framework to study membrane proteins is provided by the canonical "two-stage" model of their folding. [38,77,78] According to this model, helices are formed and inserted into the bilayer in the first stage. This is followed by specific interactions of these helices that result in higher order, functional tertiary structures. The latter step proceeds without significant modifications of the already existing elements of the secondary structure. Thus, in line with this model, functional membrane peptides have to fulfill several conditions. First, they must fold into an ordered structure. Second, they must be inserted into the membrane. Finally, they must be able to self-assemble into functional aggregates (*e.g.* channels) that exhibit specificity towards well-defined processes or substrates. For example, channels that transport both positive and negative ions, regardless of size, would have been of very limited utility to a protocell. This simple view at the emergence of membrane proteins raises a host of questions. Which amino acid sequences are capable of folding and how do membranes mediate this process? What determines the stability of the inserted, transmembrane orientation of a peptide relative to the orientation parallel to the water-membrane interface? How strong are helix-helix interactions in the membrane? How does a sequence modulate the interhelical recognition that drives the association of helices? What mechanisms are responsible for the specificity of self-assembled transmembrane peptides?

In this paper, we describe the results of computer simulations aimed at providing at least partial answers to these questions and, by doing so, help to understand better the role of peptides in the origin and early evolution of membrane-related cellular functions. After a brief description of the theoretical approach we will discuss folding of small peptides, their insertion into the membrane, interhelical recognition and selectivity of self-assembled channels. The paper closes with a summary presented in the context of the origin and early evolution of membrane proteins.

Molecular Dynamics Computer Simulations

The best approach to the computational study of biological systems at a molecular level has been the molecular dynamics method. In this section, we outline the basic aspects of this method that are common to the simulations described in this paper. Several books [4,46] provide exhaustive descriptions

of molecular dynamics, as applied to chemical and biological systems.

In molecular dynamics, Newton's equations of motion are solved numerically for all of the atoms in the system under study using an iterative procedure. From the positions, velocities and forces acting on the atoms at time t, new positions and velocities at time $t + \delta t$ can be calculated if the forces do not change appreciably during the interval δt, called "the time step". Typically, this time step is equal to 1 femtosecond (10^{-15} s). By repeating this procedure many times we obtain a time-history of the system, called a trajectory. State-of-the-art molecular dynamics trajectories for biological systems extend currently for 1 nanosecond to 1 microsecond (10^{-9} s – 10^{-6} s), which means that they require 10^6 – 10^9 time steps.

To solve Newton's equations of motion, the forces acting on the atoms in the system must be known. These forces depend, in turn, on the interactions between the atoms and are computed from a potential energy function, which describes the stretching of chemical bonds, the bending of valence angles, the rotation of dihedral angles and the electrostatic and van der Waals interactions between atoms. Much work has been invested in the construction of potential energy functions that successfully reproduce properties of water-soluble proteins, [20,64] pure membranes [12,41,88] and their mixtures with small solutes and peptides. [6,31,86] In our simulations, we took advantage of this work. For water, the TIP4P potential model was used, [49] which reproduces many of the thermodynamic properties of liquid water. The potentials that were used to describe the interactions between phospholipids have already been successfully applied in simulations of bilayer systems. [27,39,92] The proteins were described using the AMBER all-atom force field. [20]

The simulated systems consisted of a model peptide either embedded in a bilayer solvated by water or located at a model water-membrane interface. All components of the system were represented at the atomic level. The system was placed in a box — the primary simulation cell. The number of atoms in the system is typically in the 10^4–10^5 range. The corresponding cross-sectional length of the water-membrane interface varied between 4 and 6 nm. To obtain results for a macroscopic system from such simulations, the content of the primary cell was periodically replicated in space, forming an infinite lattice of identical simulation cells. This is a standard approach to removing edge effects in molecular simulations.

In many instances, considerable savings of computer time without a significant sacrifice of accuracy can be achieved by smoothly truncating interactions between two atoms in the system at a specified distance, typically

equal to 8-10 Å. However, phospholipid head groups and ionizable side chains of amino acids carry large charges, which interact non-negligibly over long distances, extending beyond the primary simulation cell. In these instances, truncation of interatomic interactions may be insufficiently accurate. The most common method for calculating long-ranged effects is called Particle-Mesh Ewald (PME) [40]. In PME, the long-range, electrostatic interactions are evaluated through the solution of a differential equation on a grid, using the Fast Fourier Transform (FFT) method. We have employed this approach in our simulations of transmembrane channels in phospholipids. [85]

Many properties of the system may be computed directly from a molecular dynamics simulation trajectory. Thermodynamic properties, such as the temperature, pressure, or membrane surface tension, can be expressed as simple averages over the series of configurations that form the trajectory. Structural quantities, such as conformational parameters characterizing the protein or its individual residues may also be computed in the same fashion. All these quantities calculated from computer simulations can be directly compared with the same quantities measured experimentally.

Often, a goal of computer simulations is to determine how the free energy of the system changes in the course of a chemical or biochemical process. These changes are directly related to the relative stabilities of different states of the system. Since free energies cannot be expressed as statistical averages of mechanical properties, special techniques are require for their evaluation. [13, 46] Two of these techniques were used in the simulations discussed in the subsequent sections.

In one approach, a series of simulations is performed, in which the system is constrained to several, overlapping ranges, or "windows", along an appropriately chosen, physical degree of freedom, often called "the reaction coordinate", $\boldsymbol{\xi}$. For example, to compute the free energy change accompanying the insertion of a peptide into the membrane, the distance between the center of mass of the peptide and the center of the lipid bilayer could be defined as such a coordinate. For each window, the probability, $\mathcal{P}(\xi)$, of finding the system at different values of the chosen coordinate is obtained. This probability defines the change of the free energy along ξ, $\Delta A(\xi)$, through the relation:

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

where k_B is the Boltzmann constant and T is the temperature of the system. $\Delta A(\xi)$ over all windows is obtained from the requirement that it must be a

continuous function of the chosen coordinate.

Another method for estimating free energy changes associated with the point mutations of a given amino acid in a peptide, or with evolution of the system along a "reaction coordinate", is based on the free energy perturbation method [104]. For point mutations, this method is implemented *via* an "alchemical transformation", in which the residues of interest in the wild type are perturbed into those of the mutant. In practice, residues are altered by modifying separately their point charges, van der Waals parameters and internal coordinates — *i.e.* shrinking or growing chemical bonds — as a function of a coupling parameter. [51] In the simulations described in this paper, creation or annihilation of non-bonded and internal parameters were carried out using the single topology approach, thus eliminating the need for defining distinct topologies for both the initial and the final states of the mutation. [72]

Folding of Simple Peptides at Water-Membrane Interfaces

Membrane peptides capable of performing cellular functions must fold into well-defined secondary structures. In fact, there are no known examples of integral membrane proteins that have largely disordered transmembrane segments. According to the "two-state" model, local folding is the first step towards self-assembly of membrane proteins into functional structures. In cells, this process proceeds at the water-membrane interface, but in laboratory experiments it can be observed also at water-micelle, water-oil or water-air interfaces. Numerous experimental studies demonstrated that short peptides can fold into highly stable structures at aqueous interfaces. [9, 16, 28, 29, 35, 43, 48, 99] A crucial, common characteristic of these interfaces is that a nonpolar phase is adjacent to water.

The ability of small peptides to organize at aqueous interfaces was examined by performing a series of large-scale, molecular dynamics computer simulations of several peptides composed of two amino acids, nonpolar L-leucine (L) and polar L-glutamine (Q) peptides differed in size and sequence of the amino acids. The simplest models studied were dipeptides LL, LQ, QL and QQ at the water-hexane interface [23, 24]. Although these peptides were too short to form secondary structures, they represented very good models

for examining conformational preferences of the peptide backbone as a function of the sequence. It was found that the preferred conformations of the backbone corresponded to optimal interactions of the side chains with the media of similar polarity. Dipeptides containing one polar (Q) and one nonpolar (L) residue adopted orientations, in which Q was immersed in water and L was buried in hexane. Dipeptides made of two polar or two nonpolar amino acids adopted orientations which maximized the exposure of their side chains to hexane and water, respectively.

The studies of L/Q peptides were extended to two heptamers, LQQLLQL and LQLQLQL, designed to maximize the interfacial stability of an α-helix and a β-strand, respectively, by exposing their polar side chains to water and their nonpolar side chains to a nonpolar phase. [24] Such structures are called amphipathic. Both amphipathic structures were highly stable at the interface. A similar result was obtained for the LQQLLQQLLQL undecamer, which is amphipathic in the α-helical conformation [24]. When this peptide was initially assigned the β-strand conformation, which is not amphipathic, it rapidly underwent several conformational transitions to adopt a nearby amphipathic (but not helical) structure. [22] During the course of the simulation a few intramolecular hydrogen bonds along the backbone were formed that are characteristic of a helical structure. All the intermediate structures were amphipathic. However, complete refolding into the expected α-helix was not observed in a 160 ns trajectory.

Finally, the transition of an undecamer, composed entirely of L-leucine residues, from a disordered structure in aqueous solution to an α-helix at the interface between water and hexane was investigated. [25] Complete folding of a peptide was accomplished for the first time in computer simulations that explicitly included the solution environment. The poly-L-leucine peptide was initially placed in water as a random coil. It rapidly translocated to the interface; in this environment, the nonpolar L side chains could be partially removed from the aqueous medium. Once at the interface, the peptide folded into a helix in 36 ns. The final structure is shown in Fig. 1. During this process, some polar groups in the backbone became dehydrated, facilitating formation of intramolecular hydrogen bonds along the backbone. The resulting structure of the peptide became more hydrophobic and partitioned further into the hexane phase. This, in turn, created a favorable environment for the emergence of additional structure-forming intramolecular hydrogen bonds. The most favorable orientation of the folded peptide was parallel to the interface. The free energy of perpendicular orientations with the N-terminus

buried in hexane was less favorable by only 4 kcal mol^{-1}. In contrast, the perpendicular orientation with C-terminus inside the hexane phase was less favorable than the parallel orientation by 12 kcal mol^{-1} because polar groups in the backbone at this end of the peptide are not involved in intramolecular hydrogen bonding but, instead, prefer to be hydrated. This result indicates that there is a preferred direction of peptide insertion into the membrane.

The simulations reveal several basic principles that govern the sequence-dependent organization of peptides at interfaces and thereby determine their ability to perform protocellular functions.

Short peptides tend to accumulate at interfaces and acquire ordered structures, provided that they have a proper sequence of polar and nonpolar residues. The specific identity of the amino acids appears to be less important for this process, a desirable protobiological property. The driving force that enables or enhances secondary structure formation for proteins interacting with or incorporated into membranes is the hydrophobic effect, which is manifested at aqueous interfaces as a tendency for polar and nonpolar groups of the solute to segregate into the aqueous and nonpolar phases, respectively. The emerging amphipathic structures are strongly favored.

Among these structures, α-helices are especially stable because they are further stabilized by intramolecular hydrogen bonding interactions. In bulk water, these interactions do not contribute to the stability of helices because of competing interactions between hydrogen bonding centers and water molecules.

If peptides consist of nonpolar residues only, they become inserted into the nonpolar phase. As demonstrated by the example of the L-leucine undecamer, nonpolar peptides tend to fold into an α–helix as they partition into the nonpolar medium. Once in the nonpolar environment, the peptides can readily change their orientation with respect to the interface from parallel to perpendicular, for example in response to local electric fields. [91] The ability of nonpolar peptides to respond to changes in external conditions may have provided a simple mechanism for transmission of signals from the environment to the interior of a protocell.

Free Energetics of Insertion of Peptides into Membranes

According to the two-state model, interfacial folding of transmembrane proteins is followed by their insertion into the bilayer. In the previous section, we have already noted one example of this phenomenon in the discussion of the poly-L-leucine undecamer. However, this example is not typical because the undecamer is not long enough to span the membrane. A transmembrane α-helical peptide must contain approximately 20 residues to extend for the full width of a membrane that is 3 nm thick. We studied one such peptide built of only two amino acids – L-leucine and L-serine (S). The peptide contains 21 residues in the sequence $(\mathrm{LSLLLSL})_3$. This sequence has been chosen such that in the α-helical form the peptide is amphipatic, *i.e.* all serine residues, which are polar, lie along the same face of the α-helix. Despite its simplicity, $(\mathrm{LSLLLSL})_3$ exhibits several interesting properties. It was shown experimentally that the peptide formed transmembrane, tetrameric ion channels in the presence of an electric field. [55] Depending on the direction of the field, the channels could transport either positive or negative ions. When the electric field was removed, the channels persisted on time scales of milliseconds before the individual peptides reverted to their resting state parallel to the water-membrane interface, indicating that the transmembrane channels do not correspond to the global free energy minimum but, instead, are weakly metastable.

Experimental results, however, provide no information about stability of individual helices in the transmembrane orientation. This will depend on the balance of hydrophobic forces, which tend to drive the nonpolar leucine residues into the membrane interior, the hydrophilic forces, which are favorable when the serine residues are located in the aqueous solution, and the interactions of unsaturated hydrogen bonding sites at both ends of the α-helix with the water. To clarify this issue we calculated the free energy of inserting the peptide into a model membrane system using a variant of the free energy window method, described briefly in the section on molecular dynamics. The resulting free energy curve is shown in Fig. 2 as a function of the location of the center-of-mass of the peptide relative to the center of the membrane ($z = 0$). The water-membrane interface is located at $z = -13$. The orientation of the peptide is dependent on the location of its center-of-mass. At $z = 0$ the peptide is approximately perpendicular to the plane of

the membrane, with an end in each of the aqueous phases. In contrast, at $z = 13$, the peptide lies parallel to the interface such that the serine residues point towards the water. The two main features of the curve in Fig. 2 are that the in-plane state is approximately 20 kcal mol^{-1} more stable then the transmembrane state and that the latter state corresponds to a broad and shallow free energy minimum. This means that the transmembrane state of the peptide is, at best, only weakly metastable and a single peptide in this state will quickly convert to the in-plane state in the absence of an electric field.

These features of the free energy surface were borne out by independent simulations with the peptide initially located in either an in-plane or transmembrane orientation. The in-plane state was stable over the course of a 10 ns simulation. The peptide backbone remained entirely α-helical and the serine residues always pointed towards the water. The transmembrane peptide adopted a variety of mixed α-helical and 3_{10}-helical arrangements. This is due to a mismatch between the length of the peptide and the width of the membrane. The peptide in the α-helical conformation is slightly longer than the width of the membrane. Converting part of the backbone to 3_{10} lengthens the helix and allows for the formation of additional, energetically favorable, serine-water hydrogen bonds. In simulations of a system with a somewhat thicker hydrophobic membrane core, the peptide remained α-helical.

A total of three trajectories were started with the peptide in the transmembrane state. On the basis of the free energy calculations it was expected that this state would not be stable and, over time, the peptide would move to the water-membrane interface. In two of the simulations, the peptide spontaneously converted from the transmembrane to the in-plane state after 7 and 9 ns, respectively. In the third simulation, however, the peptide remained transmembrane after 18 ns. This is due to the asymmetry in the ends of the peptide. As found for the undecamer of poly-Leucine, the C-terminus interacts with the water much more strongly than the N-terminus. The flat free energy curve in the region near $z = 0$ indicates that the peptide can readily diffuse towards either of the two water-membrane interfaces. In the two simulations, in which the peptide converted to the in-plane state, it initially diffused in the direction that required dehydration of the N-terminus. Since this end interacts with water relatively weakly the conversion appears to proceed quickly with little or no free energy barrier. In contrast, in the third simulations, the peptide initially diffused in the opposite direction, which would have required dehydration of the C-terminus. Considering highly fa-

vorable interaction of this terminus with water such dehydration process is unlikely. Instead, it is expected that in a sufficiently long molecular dynamics trajectory the peptide would diffuse back to the center of the bilayer and, eventually, converted to the in-plane state by dehydrating its N-terminus.

Helix Association in Membranes

Single transmembrane helices are rarely capable of performing biological functions. Instead, they form functional units after self-assembling into higher order structures. However, not all helices self-assemble. Consequently, it is necessary to understand sequence-specific interhelical recognition before we can predict the kinds of structures that could have formed in protocellular membranes.

The simplest models for peptide association are helical dimers. Although they cannot form channels, some are biologically active. Moreover, it is assumed that dimer formation is the first step in aggregation into higher order assemblies. For example, it has been suggested that tetrameric channels are formed as "dimers of dimers". [102]

Conveniently, simple experimental models have recently been developed for helix associations in membranes. [47, 59, 103] The best-studied system, both structurally and thermodynamically, is the 24-residue transmembrane region of glycophorin A (GpA). GpA forms non-covalent dimers through the reversible association of its membrane-spanning domain, [17,58] which adopts an α-helical conformation. [63] On the basis of extensive random mutagenesis investigations, a model of the GpA transmembrane dimer has been proposed, in which helix-helix association results from the specific interaction of seven residues located on one face of each α-helix. [1, 57, 59, 93] This model has been recently confirmed by NMR spectroscopy of 40-residue peptides that contained the transmembrane segment, solvated in detergent micelles. [63] This study also showed that the α-helices formed a right-handed, coil-coiled structure. This model is shown in Fig. 3

The dissociation free energy of the GpA dimer in a detergent pentaoxyethylene (C8E5) has been estimated to be about 9.0 kcal mol^{-1}. [44] It has been also demonstrated that single-residue mutations can markedly influence the free energy of association of the helices. Mutants in which either one of two leucine residues were substituted with alanine or glycine was substituted with isoleucine were found to be less stable than the wild-type

dimer by 1–3 kcal/mol. [42, 44] In the model of the GpA dimer, all these residues are involved in interhelical interactions.

To gain insight into the various contributions that govern the specific interactions of transmembrane α-helices, we simulated both the dissociation of helices and the effect of point mutations on the free energy of dissociation. [26] Any mutation that destabilizes the dimer must decrease the dissociation free energy.

In the first step, we simulated the wild-type dimer of GpA in a lamella of dodecane, placed between two lamellae of water. The width of the dodecane layer was approximately the same as that of the hydrophobic core of a palmitoyloleylphosphatidylcholine (POPC membrane. The starting structure was based on the NMR model of the dimer. [63] The comparison between the calculated, time-averaged structure of the dimer after 600 ps of molecular dynamics trajectory and the nuclear Overhauser effect data of MacKenzie *et al.* [63] provided an assessment of the accuracy of our model and the potential energy functions utilized. The distance root mean square deviation was less than 1.5 Åand the monomers remained α-helical.

Next, we separated the helices in a series of molecular dynamics simulations using the distance separating the centers of mass of the two helices as the reaction coordinate, $\boldsymbol{\xi}$. The free energy of dissociation was estimated using free energy perturbation theory. [104] The complete pathway joining the dimer at $\xi \simeq 6.5$ Å, to the dissociated state at $\xi \simeq 20$ Å, was divided into a series of intermediate states corresponding to different values of the reaction coordinate. The free energy of dissociation is given as the sum of free energy differences between consecutive states.

As can be seen in Fig. 4, the estimated free energy is approximately equal to 11.4 ± 0.3 kcal mol^{-1}. By comparison, the experimentally determined free energy of dissociation in the detergent pentaoxyethylene (C8E5) is equal to 9.0 ± 0.1 kcal mol^{-1}. [44] Since the molecular environments of GpA in the computational and experimental studies are different — *i.e.* dodecane *versus* C8E5 — the free energies are not expected to be identical, but should be similar. Given that the volume accessible to the α–helix dimer is much smaller in C8E5 than in dodecane, which corresponds to a smaller entropic contribution, we expect the association would be favored in the micellar environment. This hypothesis is supported by the observation that the free energy of dissociation is higher dodecane than C8E5.

Based on similar considerations, it is expected that the influence of the surroundings on the computed point mutations would, in principle, be lim-

ited due to compensation of entropic effects in the free energy differences. However, this may not be necessarily the case. The I76L point mutation was carried out by decoupling the annihilation of the electrostatic and the van der Waals and internal parameter contributions. Not too surprisingly, the free energy difference for the electrostatic term averaged to 0.0 kcal mol^{-1}. The contribution due to the modification of the van der Waals parameters and the participating chemical bonds and valence angles, and, therefore, the total free energy difference, is equal to 0.4 kcal mol^{-1}, which somewhat underestimates the experimental value of 1.7 kcal mol^{-1}. [44] It should be emphasized, however, that, as the L–isoleucine is transformed into L–alanine, the disruptive effect of the mutation witnessed in C8E5 should be less pronounced in dodecane, because of the greater volume accessible to the α–helices in that environment.

The free energy change upon single-point mutation is fairly small. In contrast, the free energy of dissociation is large and positive. This might suggest that the dimer is strongly favored for both the wild-type and the mutants. This would indeed be the case if we only considered the equilibrium between the associated state and separated, *transmembrane* helices. However, this comparison is not appropriate because, as we have already pointed out, the transmembrane orientation of helices is not stable. The free energy of insertion of a helical peptide into the membrane is positive (unfavorable) and may be substantial. [10, 11] Thus, the equilibrium that needs to be considered is between the transmembrane dimer and the individual helices at the water-membrane interface. This equilibrium is governed by the balance between the unfavorable free energy of insertion into the membrane and favorable free energy of interhelical association. This balance could be subtle, and modest changes in either term could shift the equilibrium, possibly disrupting dimerization.

So far, we have only discussed the possibility that peptide association follows insertion into the membrane. Alternatively, peptides might assemble at interfaces and only then become inserted into a lipid bilayer. This might seem plausible especially for amphipathic peptides because their aggregates could be stabilized by interhelical hydrogen bonds between polar residues. Then, nonpolar residues would be exposed to the environment, which should promote insertion into the nonpolar membrane. To test this mechanism of aggregation, we have carried out simulation studies on dimers of the $(LSLLLSL)_3$ peptide, which has already been discussed in the previous section. When the dimer was placed parallel to the plane of the membrane at

the water-membrane interface, it dissociated in less than 2 ns. The interfacial water molecules successfully competed for the serine hydrogen bonding sites, which led to the loss of serine-serine hydrogen bonds. Additionally, interactions between electrical dipoles associated with the α-helices are highly unfavorable in a parallel arrangement. These results indicate that self-assembly of peptides at interfaces is unlikely. In contrast, a transmembrane dimer was found to be stable over the course of a 15 ns simulation. Near the ends of the dimer, serine-serine hydrogen bonds were lost in favor of water-serine hydrogen bonds, allowing water molecules to penetrate the membrane around the peptide. This, in turn, might increase rates of non-specific permeation of ions and polar solutes across membranes. [33, 70, 71, 76, 98] However, the serine-serine hydrogen bonds in the middle of the dimer remained intact, keeping the helices in the dimer together. These results confirm that association of two peptides appears to increase the stability of the transmembrane state relative to isolated monomers.

Simulation of a Model Transmembrane Proton Transport System

Aggregates of membrane proteins are of special interest if they can perform important cellular functions. One such function is transport of protons across membranes, which is an essential process for both bioenergetics of modern cells and the origins of cellular life. All living systems convert environmental energy into chemical energy by using transmembrane proton gradients to drive the synthesis of adenosine triphosphate (ATP) from adenosine diphosphate (ADP). ATP, in turn, is used as a source of energy to drive many cellular reactions. The ubiquity of this process in biology suggests that even the earliest cellular systems relied on proton gradients to harvest the energy needed for their survival and growth from the environment. In contemporary cells, proton transfer is assisted by large, complex proteins embedded in membranes. Could the same process have been accomplished with the aid of similar, but much simpler peptides that could have existed in the protobiological milieu?

To answer this question it is desirable to have a protein model which is small, has a well known structural motif, yet which operates with the efficiency and control of more complex proteins. This led us to study the

Influenza-A M_2 protein, which forms small, voltage-gated proton channels. [45,60,67,73,74,82,95,101] The M_2 protein contains 97 amino acids, including a single transmembrane domain 19 residues long. Not all residues, however, are essential for transport. Active channels have been reconstituted from a synthetic peptide containing a subset of only 25 amino acids, including the transmembrane region, with no loss in specificity or efficiency. [36]

In lipid bilayers, four identical protein fragments, each folded into an α-helix, aggregate to form small channels spanning the membrane. Protons are conducted through a narrow pore in the middle of the channel. Compared with a well-studied, proton-permeating peptide, gramicidin A, the rate of proton transport across the truncated M_2 channel is over 1000-fold faster. Remarkably, in contrast to gramicidin A, the M_2 channel is virtually impermeable to alkali ions, such as Na^+ and K^+. This combination of efficiency and specificity makes M_2 an excellent, simple model to study the formation of proton gradients across membranes.

The channel is large enough to contain water molecules and is normally filled with water. In analogy to the mechanism of proton transfer in some other channels, [3, 84] it has been postulated that protons are translocated along the network of properly aligned water molecules filling the pore. This mechanism, however, must involve an additional, important step because the channel contains four L-histidine (H) amino acid residues, one from each of the helices, which are sufficiently large to occlude the pore and interrupt the water network. The L-histidine residues have been implicated in gating protons. Due to their size, they ensure channel selectivity by blocking small ions, such as Na^+ and K^+, from permeating the membrane but provide a mechanism for proton transport. The role of the L-histidines in gating is supported by findings that point mutations, in which the L-histidines are substituted by other residues greatly impede the ability of M_2 to transport protons.

Two mechanisms of gating have been proposed, which rely on the ability of each L-histidine to become positively charged by accepting an additional proton. In one mechanism, all four L-histidines acquire a proton and, due to repulsion between their positive charges, move away from one another, thus opening the channel. [83] The alternative mechanism involves the ability of protons to move between different atoms in a molecule (tautomerization). In this mechanism, a proton is captured on one side of the gate while a second proton is released from the opposite side, and the molecule returns to the initial state through tautomerization. [73] These two mechanisms are shown

schematically in Fig. 5.

Atomic-level molecular dynamics simulations were designed to test these two mechanisms. [75, 85] The model system used in the study contained a bilayer membrane made of phospholipid, dimyristoylphosphatidylcholine (DMPC), which is a good model of the biological membranes forming cellular boundaries. Both sides of the bilayer were surrounded by water which simulated the environment inside and outside the cell. Embedded in the membrane was a channel made of 25 amino acids fragments of the Influenza-A M_2 protein, and enough sodium counterions to maintain system neutrality. Several protonation states of L-histidine residues were considered. They represented different intermediate states of the channel predicted by the two proposed mechanisms of proton transport. The simulations revealed that all intermediate states of the system involved in the tautomerization mechanism were structurally stable and the arrangement of water molecules in the channel was conducive to the proton transport. In contrast, in the four-protonated state, postulated to exist in the gate-opening mechanism, the electrostatic repulsion between the L-histidine residues appeared to be so large that the channel lost its structural integrity and one helix moved away from the remaining three. These results indicate that translocation along a network of water molecules in the channel and tautomerization of the L-histidine residues is a likely mechanism of proton transport whereas a mechanism involving protonation of all four L-histidines is unlikely. A possibility of gate opening after protonation of two rather than four L-histidines has not been excluded.

These results not only explain how a simple protein system can achieve highly efficient and selective passive proton transport (i.e. transport along the concentration gradient) across cell walls, but also how the system can be genetically re-engineered to become a simple directional, reversible proton pump. First, M_2 must be coupled with a chromophore capable of releasing a proton in response to light. Several very simple chromophores, such as polycyclic aromatic hydrocarbons, are already known. In fact, some of them have been shown to dissolve in membranes and generate transient, light-induced proton gradients. [32] To maintain the proton gradient, it must ensured that release of the pumped proton is followed by reprotonation of the chromophore with a proton from the opposite side of the membrane. This will involve manipulating the sequence of amino acids along the pore. Cysteine scanning mutagenesis has already shown that the replacement of the pore-lining, but not other residues can modify the properties of the channel. [87] Other de-

signs are possible, based on coupling electron and proton transfer using iron or quinones. These have been recently shown to be of possible protobiological relevance. [14] If such an experimental effort were to be successful, it would demonstrate that protein-based proton pumps could have emerged early in protobiological evolution. Furthermore, such a pump could be used to provide energy to laboratory-built models of protocells and cell-like structures built for biotechnological applications. [53]

Summary

Many proteins that perform essential cellular functions are embedded in membranes that encapsulate cells or cellular components. These proteins or protein complexes are among the largest macromolecular structures found in cells and their mode of action is often complicated and subtle. This appears to create a serious difficulty from the origin of life point of view. If the functions performed by membrane proteins are essential to the existence of even the simplest cells, how could they have been performed, even less efficiently or selectively, by much simpler peptides?

In this paper we have argued that the emergence of integral membrane proteins may have been quite feasible. In fact, this may be much easier to envision than the emergence of water-soluble proteins. We have supported our arguments with results of our molecular dynamics computer simulations and a considerable body of evidence from other experimental and theoretical studies. The prerequisite for the formation of functional membrane proteins was the existence of peptides containing 20–25 amino acid, which were sufficiently long to span a membrane. This length requirement is rather modest, considering that functional water-soluble proteins need to be markedly larger. For example, the shortest, protobiologically relevant proteins contain 45 amino acids [90] and the simplest self-reproducing protein system consists of proteins built of 33 amino acids. [56]

Many peptides are attracted to water-membrane or water-oil interfaces. Once at the interface, most nonpolar peptides spontaneously fold into α-helices. Peptides that contain both polar and nonpolar amino acids tend to adopt amphipathic structures, in which amino acid side chains are immersed in media of similar polarity. Whenever the sequence permits, peptides fold into amphipathic helices at interfaces. The formation of ordered, helical structures is primarily governed by the sequence of polar and nonpolar amino

acids. Considering that specific identities of side chains is less important, the existence of helical peptides in interfacial, protocellular environments should not have been rare.

Helical peptides located parallel to the interface could insert into the membrane and adopt a transmembrane conformation. However, insertion of a single helix usually involves a positive free energy change, even for fully nonpolar peptides. The main reason why insertion is unfavorable is that polar groups in the peptide backbone and some side chains, which remain at least partially hydrated in water, become completely desolvated. The loss of solvation free energy is smaller for helices than for disordered structures because polar groups in the backbone are involved in intramolecular hydrogen bonding.

The unfavorable free energy of insertion can be regained by spontaneous association of peptides in the membrane into homomeric or heteromeric multimers. The first step in this process is the formation of dimers, although the most common structures involve aggregates of 4–7 helices. The helices could readily arrange themselves such that they form pores capable of transporting ions and small molecules across membranes. The stability of transmembrane aggregates of simple proteins is often marginal and, therefore, it can be regulated by environmental conditions, such as external electric fields or the specific nature of phospholipid headgroups, [15,19,54,91] or by small changes in the sequence of amino acids. [42,44] This ability to respond to environmental signals might have led to the earliest, although quite imprecise, regulation of transmembrane functions.

Clearly, a key step in the earliest evolution of integral membrane proteins was the emergence of selectivity for specific substrates. The selectivity of early channels was determined to some extent by all residues lining their lumen, which interact with substrates *via* electrostatic and van der Waals interactions. [79] However, many contemporary simple channels employ filters or gates as the primary way to achieve selectivity. [68,73,80] From the evolutionary standpoint it is a very convenient solution because it requires placing just one or only a very few properly chosen residues in certain positions along the channel rather than imposing conditions on the whole sequence.

Many additional steps were required before simple aggregates of transmembrane peptides reached the structural and functional complexity, diversity and refinement of contemporary integral membrane proteins. The helices were connected by extra-membrane, hydrophilic linkers to stabilize them inside the membrane. The resulting, large proteins aggregated to even larger,

higher-order structures. In many instances this step involved gene duplication. Protein sequences became optimized for highly specific functions. Perhaps most importantly, membrane proteins acquired large, water-soluble domains, which play a regulatory role or help to supply energy for active transport. This more advanced evolution of membrane proteins has been a subject of extensive studies. [77] In the process, some intriguing connections between ion channels and enzymes have been uncovered. [62] The evolutionary history of membrane proteins is special interest because it opened the doors for the emergence of multicellular organisms endowed with nervous systems.

Acknowledgments

This work was supported by the grants from the NASA Exobiology Program and from the NASA Astrobiology Institute.

References

[1] Adams, P. D., D. M. Engelman, and A. T. Brünger: 1996, 'Improved prediction for the structure of a dimeric transmembrane domain of glycophorin A obtained through global searching'. *Proteins: Structure, Function and Genetics* **26**, 257–261.

[2] Åkerfeldt, K., J. Lear, Z. Wasserman, L. Chung, and W. DeGrado: 1993, 'Synthetic peptides as models for ion channel proteins'. *Acc. Chem. Res.* **26**, 191–197.

[3] Akeson, M. and D. Deamer: 1990, 'Proton conductance by the Gramicidin water wire — model for proton conductance in the F_1F_0 ATPases'. *Biochim. Biophys. Acta* **60**, 101–109.

[4] Allen, M. and D. Tildesley: 1987, *Computer Simulation of Liquids.* Oxford: Oxford University Press.

[5] Apel, C., D. Deamer, and M. Mautner: 2002, 'Self-assembled vesicles of monocarboxylic acids and alcohols: conditions for stability and for the encapsulation of biopolymers'. *Biochim. Biophys. Actai - Biomembranes* **1559**, 1–9.

[6] Bassolino-Klimas, D., H. Alper, and T. Stouch: 1995, 'Mechanism of solute diffusion through lipid bilayer membranes by molecular dynamics simulation'. *J. Am. Chem. Soc.* **117**, 4118–4129.

[7] Bayley, H.: 1999, 'Designed membrane channels and pores'. *Curr. Opin. Biotech.* **10**, 94–103.

[8] Bechinger, B.: 2000, 'Understanding peptide interactions with the lipid bilayer: a guide to membrane protein engineering'. *Curr. Opin. Chem. Biol.* **4**, 639–644.

[9] Bechinger, B., M. Zasloff, and S. Opella: 1993, 'Structure and orientation of the antibiotic peptide magainin in membranes bu solid-state nuclear-magnetic-resonance spectroscopy'. *Prot. Sci.* **2**, 2077–2084.

[10] Ben-Shaul, A., N. Ben-Tal, and B. Honig: 1996, 'Statistical thermodynamic analysis of peptide and protein insertion into lipid membranes'. *Biophys. J.* **71**, 130–137.

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

[12] Berger, O., O. Edholm, and F. Jähnig: 1997, 'Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature'. *Biophys. J.* **72**, 2002–2013.

[13] Berne, B. and J. Straub: 1997, 'Novel methods of sampling phase space in the simulation of biological systems'. *Curr. Opin. Struct. Biol.* **7**, 181–189.

[14] Bernstein, M., J. Dworkin, S. Sandford, and L. Allamandola: 2001, 'Ultraviolet Irradiation of Naphthalene in H2O Ice: Implications for Meteorites and Biogenesis'. *Meteoritics and Planetary Science* **36**, 351–358.

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

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

[17] Bormann, B. J., W. J. Knowles, and V. T. Marchesi: 1989, 'Synthetic peptides mimic the assembly of transmembrane glycoproteins'. *J. Biol. Chem.* **264**, 4033–4037.

[18] Boyd, D., C. Schierle, and J. Beckwith: 1998, 'How many membrane proteins are there'. *Protein Sci.* **7**, 201–205.

[19] Cafiso, D.: 1994, 'Alamethicin: A peptide model for voltage gating and protein-membrane interactions'. *Ann. Rev. Biophys. Biomol. Struct.* **23**, 141–165.

[20] Case, D., D. Pearlman, J. Caldwell, T. Cheatham, III, W. Ross, C. Simmerling, T. Darden, K. Merz, R. Stanton, A. Cheng, J. Vincent, M. Crowley, V. Tsui, R. Radmer, Y. Duan, J. Pitera, I. Massova, G. Siebel, U. Singh, P. Wiener, and P. Kollman: 1999, 'AMBER 6'. University of California, San Francisco.

[21] Chang, S.: 1993, 'Prebiotic synthesis in planetary environments'. In: J. M. Greenberg (ed.): *The Chemistry of Life's Origins.* Amsterdam: Kluwer, pp. 259–299.

[22] Chipot, C., B. Maigreit, and A. Pohorille: 1999, 'Early events in the folding of an amphipathic peptide at the water-hexane interface. A multi-nanosecond molecular dynamics study'. *Proteins Struct. Func. Genet.* **36**, 383–399.

[23] Chipot, C. and A. Pohorille: 1997a, 'Conformational Equilibria of Terminally Blocked Single Amino Acids at the Water-Hexane Interface. A Molecular Dynamics Study'. *J. Phys. Chem. B* **102**, 281–290.

[24] Chipot, C. and A. Pohorille: 1997b, 'Structure and Dynamics of Small Peptides at Aqueous Interfaces. A Multi-Nanosecond Molecular Dynamics Study.'. *J. Mol. Struct. (THEOCHEM)* **398**, 529–535.

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

[26] Chipot, C. and A. Pohorille: 2002. Unpublished results.

[27] Chiu, S.-W., M. Clark, V. Balaji, S. Subramaniam, H. L. Scott, and E. Jakobsson: 1995, 'Incorporation of surface tension into molecular dynamics simulation of an interface: A fluid phase lipid bilayer membrane'. *Biophys. J.* **69**, 1230–1245.

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

[29] Cornut, I., B. Desbat, J. Turlet, and J. Dufourcq: 1996, '*In situ* study by polarization modulated Fourier transform infrared spectroscopy of the structure and orientation of lipids and amphipathic peptides at the air-water interface'. *Biophys. J.* **70**, 305–312.

[30] Cronin, J. and S. Chang: 1993, 'Organic matter in meteorites: Molecular and isotopic analyses of the Murcheson meteorite'. In: J. M. Greenberg (ed.): *The Chemistry of Life's Origins.* Amsterdam: Kluwer, pp. 200–258.

[31] Damodaran, K., K. Merz, Jr., and B. Gaber: 1995, 'Interaction of small peptides with lipid bilayers'. *Biophys. J.* **69**, 1299–1308.

[32] Deamer, D.: 1992, 'Polycyclic aromatic hydrocarbons: primitive pigment systems in the prebiotic environment'. *Advances in Space Research* **12**, 183–189.

[33] Deamer, D. and J. Nichols: 1989, 'Proton flux mechanisms in model and biological membranes'. *J. Membrane Biol.* **107**, 91–103.

[34] Deamer, D. W. and R. M. Pashley: 1989, 'Amphiphilic compomenets of the Murchison Carbonaceous Chondrite: Surface Properties and Membrane Formation'. *Origins Life Evol. Biosphere* **19**, 21–38.

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

[36] Duff, K. C. and R. H. Ashley: 1992, 'The Transmembrane Domain of Influenza A M_2 Protein forms Amantidine Sensitive Proton Channels in Planar Lipid Bilayers'. *Virology* **190**, 485–489.

[37] Dworkin, J., D. Deamer, S. Sandford, and L. Allamandola: 2001, 'Self-assemblying amphiphilic molecules: Synthesis in simulated interstellar/precometary ices'. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 815–819.

[38] Engelman, D. and T. Steitz: 1981, 'The spontaneous insertion of proteins into and across membranes: The helical hairpin hypothesis'. *Cell* **23**, 411–422.

[39] Essmann, U., L. Perera, and M. Berkowitz: 1995a, 'The origin of the hydration interaction of lipid bilayers from MD simulation of dipalmitoylphosphatidylcholine membranes in gel and liquid crystalline phases.'. *Langmuir* **11**, 4519–4531.

[40] Essmann, U., L. Perera, M. Berkowitz, T. Darden, H. Lee, and L. Pedersen: 1995b, 'A smooth particle mesh Ewald method'. *J. Chem. Phys.* **103**(19), 8577–8593.

[41] Feller, S. E.: 2000, 'Molecular dynamics simulations of lipid bilayers'. *Curr. Opin. Colloid Interface Sci.* **5**, 217–223.

[42] Fisher, L. E., D. M. Engelman, and J. N. Sturgis: 1999, 'Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain'. *J. Mol. Biol.* **293**, 639–651.

[43] Flach, C., J. Brauner, J. Taylor, R. Baldwin, and R. Mendelsohn: 1994, 'External reflection FTIR of peptide monolayer films in situ at the air/water interface: Experimental design, spectra-structure correlations and effects of hydrogen-deuterium exchange'. *Biophys. J.* **67**, 402–410.

[44] Fleming, K. G., A. L. Ackerman, and D. M. Engelman: 1997, 'The effect of point mutations on the free energy of transmembrane α-helix dimerization'. *J. Mol. Biol.* **272**, 266–275.

[45] Forrest, L. and M. Sansom: 2000, 'Membrane simulations: bigger and better?'. *Curr. Opin. Struct. Bio.* **10**, 174–181.

[46] Frenkel, D. and B. Smit: 1986, *Understanding Molecular Simulations.* San Diego: Academic Press.

[47] Gratowski, H., J. D. Lear, and W. F. DeGrado: 2001, 'Polar side chains drive the association of model transmembrane peptides'. *Proc. Natl. Acad. Sci. USA* **98**, 880–885.

[48] Ishiguro, R., N. Kimura, and S. Takahashi: 1993, 'Orientation of fusion-active synthetic peptides in phospholipid bilayers: Determination by Fourier-transform infrared-spectroscopy'. *Biochem.* **32**, 9792–9797.

[49] Jorgensen, W., J. Chandrasekhar, J. Madura, R. Impey, and M. Klein: 1983, 'Comparison of simple potential functions for simulating liquid water'. *J. Chem. Phys.* **79**, 926–935.

[50] Keire, D. and F. T.G.: 1996, 'The conformation of substance P in lipid environments'. *Biophys. J.* **70**, 1716–1727.

[51] Kollman, P.: 1993, 'Free Energy Calculations — Applications to Chemical and Biochemical Phenomena'. *Chem. Rev.* **93**, 2395.

[52] Kraulis, P.: 1991, 'MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures'. *J. App. Crys.* **24**, 946–950.

[53] Lanyi, J. K. and A. Pohorille: 2001, 'Proton pumps: mechanism of action and applications'. *Trends Biotechol.* **19**, 140–144.

[54] Lear, J., J. Schneider, P. Kienker, and W. DeGrado: 1997, 'Electrostatic effects on ion selectivity and rectification in designed ion channel peptides'. *J. Am. Chem. Soc.* **119**, 3212–3217.

[55] Lear, J., Z. Wasserman, and W. DeGrado: 1988, 'Synthetic amphiphilic peptide models for protein ion channels'. *Science* **240**, 1177–1181.

[56] Lee, D. H., K. Severin, Y. Yokobayashi, and M. Ghadiri: 1997, 'Emergence of symbiosis in peptide self-replication through a hypercyclic network'. *Nature* **390**(6660), 591–594.

[57] Lemmon, M. A. and D. M. Engelman: 1987, 'Specificity and promiscuity in membrane helix interactions'. *Quarterly Reviews of Biophysics.* **27**, 157–218.

[58] Lemmon, M. A., J. Flanagan, J. Hunt, B. Adair, B. J. Bormann, C. Dempsey, and D. M. Engelman: 1992a, 'Glycophorin A dimerization is driven by specific interactions between transmembrane α-helices'. *J. Biol. Chem.* **267**, 7683–7689.

[59] Lemmon, M. A., J. Flanagan, H. R. Treutlein, J. Zhang, and D. M. Engelman: 1992b, 'Sequence specificty in the dimerization of transmembrane α-helices'. *Biochemistry* **31**, 12719–12725.

[60] Lin, T. I. and C. Schroeder: 2001, 'Definitive assignment of proton selectivity and attoampere unitary current to the M_2 ion channel protein of influenza A virus'. *J. Virol.* **75**, 3647–3656.

[61] Luisi, P. L., P. Walde, M. Blocher, and D. J. Liu: 2000, 'Research on the origin of life: Membrane-assisted polycondensations of amino acids and peptides'. *Chimia* **54**, 52–53.

[62] L.Y., J. and J. Y.N.: 1992, 'Tracing the roots of ion channels'. *Cell* **69**, 715–718.

[63] MacKenzie, K. R., J. H. Prestegard, and D. M. Engelman: 1997, 'A transmembrane helix dimer: Structure and implications'. *Science* **276**, 131–133.

[64] MacKerell, Jr., A., B. Brooks, C. Brooks, III, L. Nilsson, B. Roux, and M. Won, Y.and Karplus: 1998, 'CHARMM: The Energy Function and Its Parameterization with an Overview of the Program'. In: *et al.*. Schleyer, P.v.R. (ed.): *The Encyclopedia of Computational Chemistry*, Vol. 1. New York: John Wiley and Sons, pp. 271–277.

[65] Miller, S.: 1953, 'A production of amino acids under under possible primitive earth conditions'. *Science* **117**, 528–529.

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

[67] Mould, J., J. E. Drury, S. M. Frings, U. B. Kaupp, A. Pekosz, R. A. Lamb, and L. H. Pinto: 2000, 'Permeation and activation of the M_2 ion channel of influenza A virus'. *J. Biol. Chem.* **275**, 31038–31050.

[68] Murata, K., K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. Heymannh, A. Engel, and Y. Fujiyoshi: 2001, 'Structural determinants of water permeation through aquaporin-1'. *Nature* **407**, 599–605.

[69] Oblatt-Montal, M., L. Buhler, T. Iwamoto, J. Tomich, and M. Montal: 1993, 'Synthetic peptides and four-helix bundle proteins as model

systems for the pore-forming structure of channel proteins. I. Transmembrane segment M2 of the nicotinic cholinergic receptor channel is a key pore-lining structure'. *J. Biol. Chem.* **268**, 14601–14607.

[70] Paula, S., A. Volkov, and D. Deamer: 1998, 'Permeation of Halide Anions through Phospholipid Bilayers Occurs by the Solubility-Diffusion Mechanism'. *Biophys. J.* **74**, 319–327.

[71] Paula, S., A. Volkov, A. Van Hoeck, T. Haines, and D. W. Deamer: 1996, 'Permeation of protons, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness'. *Biophys. J.* **70**, 339–348.

[72] Pearlman, D.: 1994, 'A Comparison of Alternative Approaches to Free Energy Calculations'. *J. Phys. Chem.* **98**, 1487–1493.

[73] Pinto, L., G. Dieckmann, C. Gandhi, C. Papworth, J. Braman, M. Shaughnessy, J. Lear, R. Lamb, and W. DeGrado: 1997, 'A Functionally Defined Model for the M_2 Proton Channel of Influenza A Virus Suggests a Mechanism for its Ion Selectivity'. *Proc. Natl. Acad. Sci. USA* **94**, 11301–11306.

[74] Pinto, L., L. Holsinger, and R. Lamb: 1992, 'Influenza Virus M_2 Protein had Ion Channel Activity.'. *Cell* **69**, 517–528.

[75] Pohorille, A., M. Wilson, C. Chipot, N. M.H., and K. Schweighofer: 1999, 'Interactions of small molecules and peptides with membranes'. In: J. Lesczynski (ed.): *Computational Molecular Biology*, Theoretical and Computational Chemistry. Amsterdam: Elsevier, pp. 485–526.

[76] Pohorille, A. and M. A. Wilson: 2001, 'Unassisted and assisted ion transport across membranes: Insights from computer simulations'. *Cell. Mol. Biol. Lett.* **6**, 369–374.

[77] Popot, J. L. and D. M. Engelman: 2000, 'Helical membrane protein folding, stability, and evolution'. *Annu. Rev. Biochem.* **69**, 881–922.

[78] Popot, J. L., D. M. Engelman, O. Gurel, and G. Zaccai: 1990, 'Tertiary structure of bacteriorhodopsin. Positions and orientations of helices A and B in the structural map determined by neutron diffraction'. *J. Mol. Biol.* **210**, 829–847.

[79] Roux, B., S. Berneche, and W. Im: 2000, 'Ion channels, permeation, and electrostatics: Insight into the function of KcsA'. *Biochemistry* **44**, 13295–13306.

[80] Roux, B. and R. MacKinnon: 1999, 'The cavity and pore helices the KcsA K^+ channel: Electrostatic stabilization of monovalent cations'. *Science* **424**, 100–102.

[81] Saghatelian, A., Y. Yokobayashi, K. Soltani, and M. R. Ghadiri: 2001, 'A chiroselective peptide replicator'. *Nature* **409**, 797–801.

[82] Sakaguchi, T., L. Quiang, L. Pinto, and R. Lamb: 1997, 'The Active Oligomeric State of the Minimalistic Influenza Virus M_2 Ion Channel is a Tetramer'. *Proc. Natl. Acad. Sci. USA* **94**, 5000–5005.

[83] Sansom, M., I. Kerr, and H. Son: 1997, 'The Influenza A Virus M_2 Ion Channel: A Molecular Modeling and Simulation Study.'. *Virology* **233**, 163–173.

[84] Schumaker, M. F., R. Pomes, and B. Roux: 2001, 'Framework model for single proton conduction through gramicidin'. *Biophys. J.* **80**, 12–30.

[85] Schweighofer, K. and A. Pohorille: 2000, 'Computer simulation of ion channel gating: The M_2 channel of influenza A virus in a lipid bilayer'. *Biophys. J.* **78**, 150–163.

[86] Shen, L., D. Bassolino-Klimas, and T. Stouch: 1997, 'Transmembrane helix structure, dynamics, and interactions: Multi–nanosecond, molecular dynamics simulations'. *Biophys. J.* **73**, 3–20.

[87] Shuck, K., R. A. Lamb, and L. H. Pinto: 2000, 'Analysis of the pore structure of the influenza A virus M_2 ion channel by the substituted-cysteine accessibility method'. *J. Virol.* **74**, 7755–7761.

[88] Smondyrev, A. M. and M. L. Berkowitz: 1999, 'Molecular dynamics simulation of DPPC bilayer in DMSO'. *Biophys. J.* **76**, 2472–2478.

[89] Song, L., M. Hobaugh, C. Shustak, S. Cheley, and H. Bayley: 1996, 'Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore'. *Science* **274**, 1859–1865.

[90] Szostak, J. W., P. L. Luisi, and D. P. Bartel: 2001, 'Synthesizing life'. *Nature* **409**, 387–390.

[91] Tieleman, D., H. Berendsen, and M. Sansom: 2001, 'Voltage-dependent insertion of alamethicin at phospholipid/water and octane/water interfaces'. *Biophys. J.* **80**, 331–346.

[92] Tieleman, D., S. Marrink, and H. Berendsen: 1997, 'A computer perspective of membranes: Molecular dynamics studies of lipid bilayer systems'. *Biochim. Biophys. Acta* **1331**, 235–270.

[93] Treutlein, H. R., M. A. Lemmon, D. M. Engelman, and A. T. Brünger: 1992, 'The glycophorin A transmembrane domain dimer: Sequence-specific propensity for a right-handed supercoil of helices'. *Biochemistry* **31**, 12726–12732.

[94] Wallin, E. and G. von Heijne: 1998, 'Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms'. *Protein Sci.* **7**, 1029–1038.

[95] Wang, C., K. Takeuchi, L. Pinto, and R. Lamb: 1993, 'Ion channel activity of Influenza A virus M_2 protein — Characterization of the amantadine block'. *J. Virology* **67**, 5585–5594.

[96] Weber, A. L.: 1998, 'Prebiotic amino acid thioester synthesis: Thiol-dependent amino acid synthesis from formose substrates (formaldehyde and glycolaldehyde) and ammonia'. *Orig. Life. Evol. Biosph.* **28**, 259–270.

[97] Weber, A. L.: 2001, 'The sugar model: Catalysis by amines and amino acid products'. *Orig. Life. Evol. Biosph.* **31**, 71–86.

[98] Wilson, M. and A. Pohorille: 1996, 'Mechanism of Unassisted Ion Transport across Membrane Bilayers'. *J. Am. Chem. Soc.* **118**, 6580–6587.

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

[100] Yao, S., I. Ghosh, R. Zutshi, and J. Chmielewski: 1998, 'Selective amplification by auto- and cross-catalysis in a replicating peptide system'. *Nature* **396**, 447–450.

[101] Zhong, Q., T. Hisslein, P. Moore, D. Newns, P. Pattnaik, and M. Klein: 1998a, 'The M2 channel of influenza A virus: A molecular dynamics study'. *FEBS Letters* **434**, 265–271.

[102] Zhong, Q., Q. Jiang, P. Moore, D. Newns, and M. Klein: 1998b, 'Molecular Dynamics Simulation of a Synthetic Ion Channel'. *Biophys. J.* **74**, 3–10.

[103] Zhou, F. X., H. J. Merianos, A. T. Brünger, and D. M. Engelman: 2001, 'Polar residues drive association of polyleucine transmembrane helices'. *Proc. Natl. Acad. Sci. USA* **98**, 2250–2255.

[104] Zwanzig, R.: 1954, 'High temperature equation of state by a perturbation method. I Nonpolar gasses'. *J. Chem. Phys.* **22**, 1420–1426.

Figure Captions

Figure 1. Snapshot of the α-helical poly-L-leucine at the water-membrane interface. The leucine side chains are shown in black, the backbone and blocking groups in dark gray, the membrane in light gray, and the water molecules in medium gray and white.

Figure 2. Free energy profile of the center-of-mass of the $(LSLLLSL)^3$ peptide as a function of its position relative to the center of the model membrane. The center of the membrane is located at $z = 0$ and the water membrane interface is located at $z = 15$.

Figure 3. Dimeric transmembrane region of glycophorin A (GpA). The two alpha-helices formed by residues 73 to 95, blocked by Ace- and -NHMe groups, interact with a right-handed crossing angle of -40 degrees. Residue Ile 76, involved in the alchemical transformation described in the text, is highlighted with a transparent CPK representation. The image was generated using MOLSCRIPT [52]

Figure 4. Free energy profile of GpA dissociation. The minimum at $z = 5.7$ corresponds to close contact of the two helices.

Figure 5. Schematics of two proposed models of gated proton transport in M2: (a) a proton shuttle mechanism mediated by the His^{37} residues, where the initial state is regenerated by tautomerization and (b) a structural model where the gating is mediated by steric forces resulting from protonation of one or more His^{37} residues.

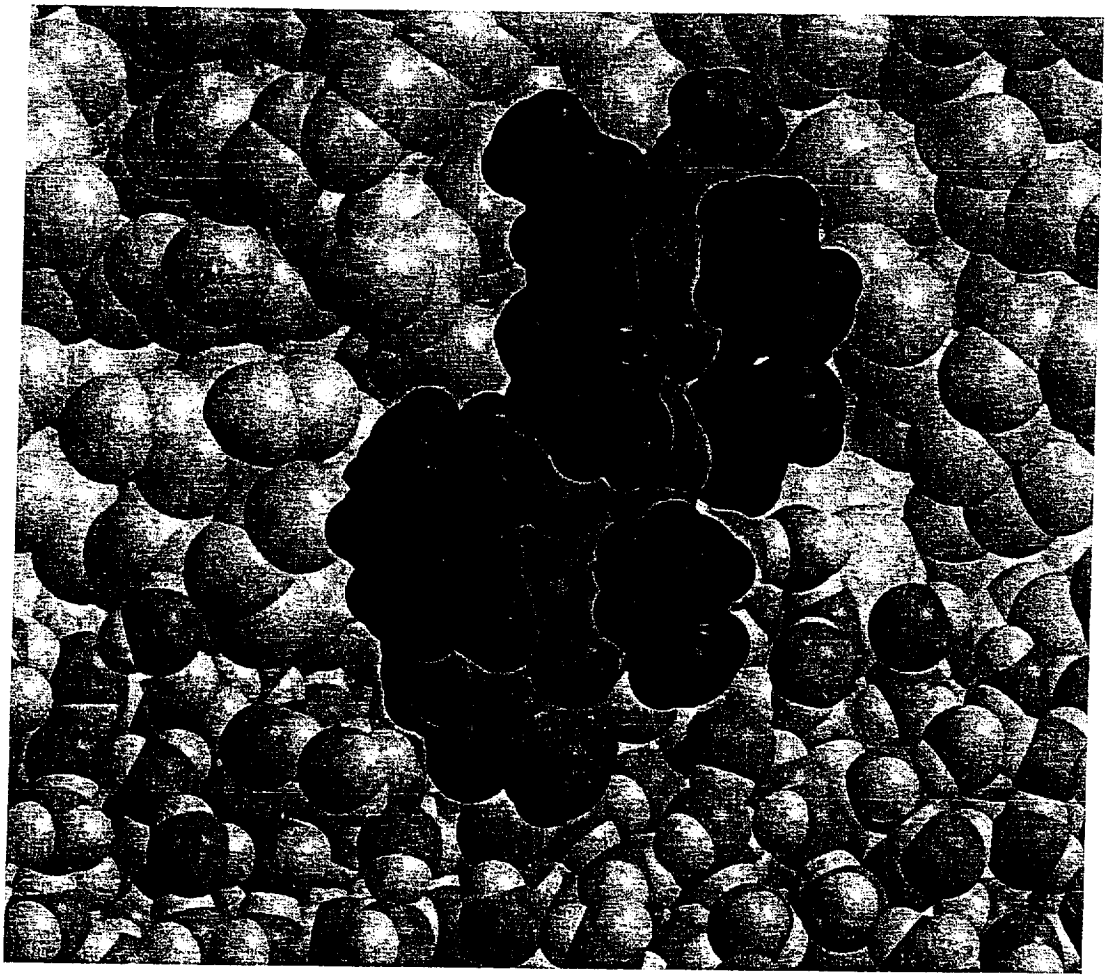

Figure 1:

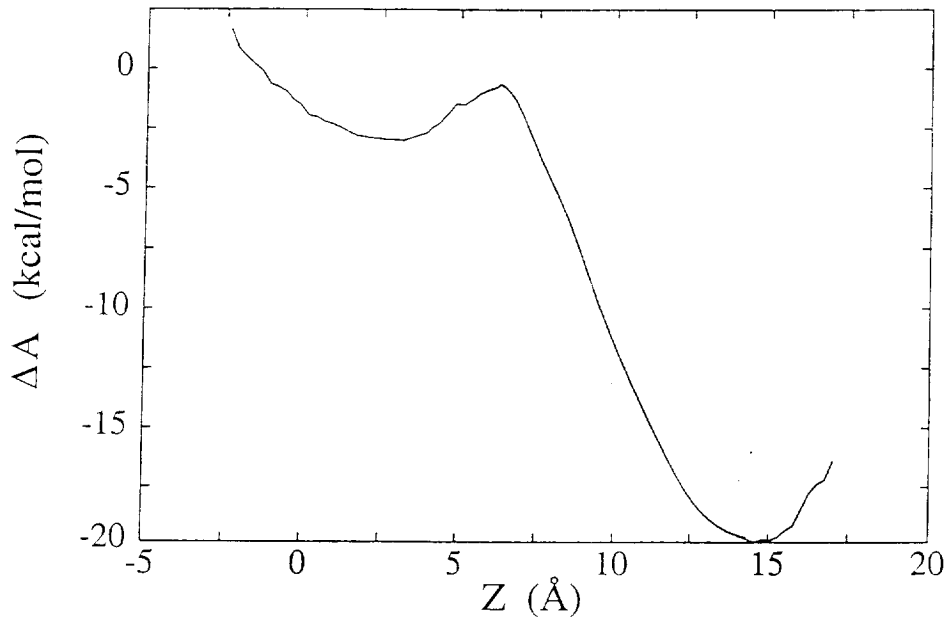

Figure 2:

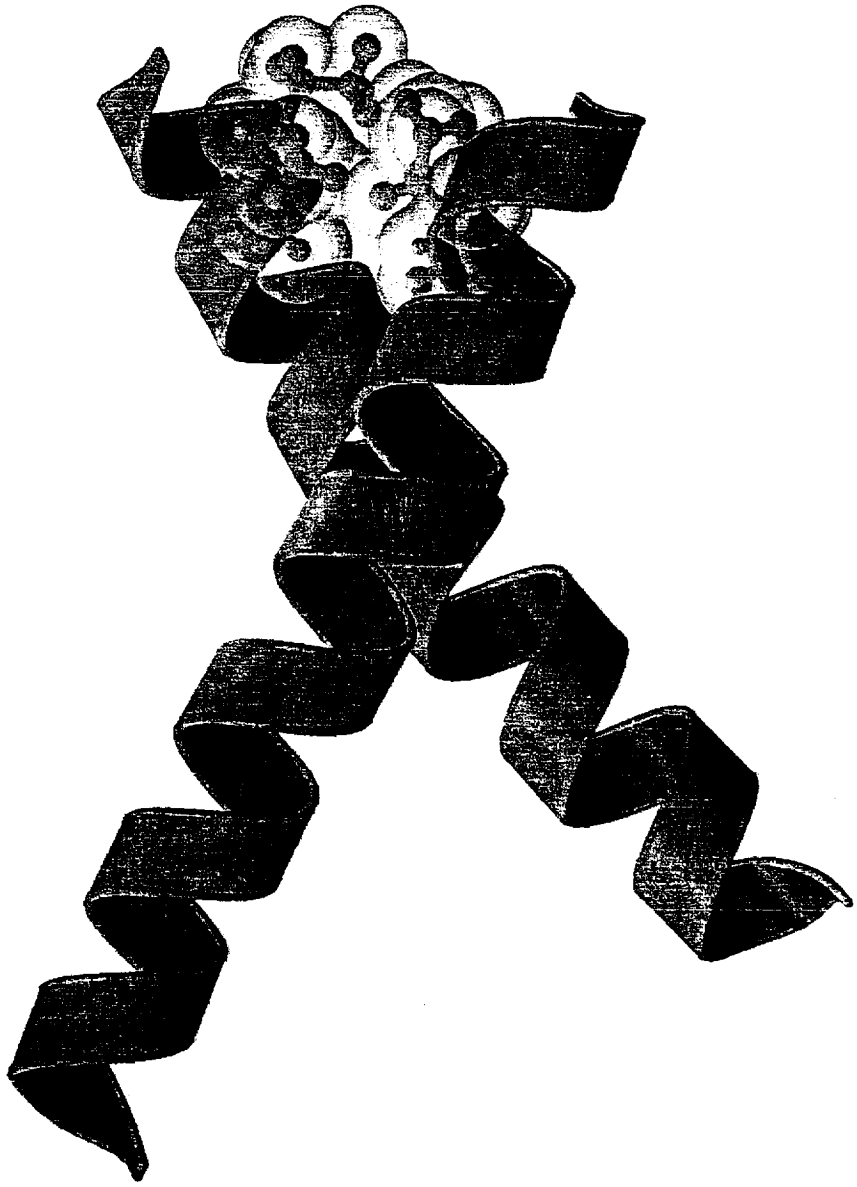

Figure 3:

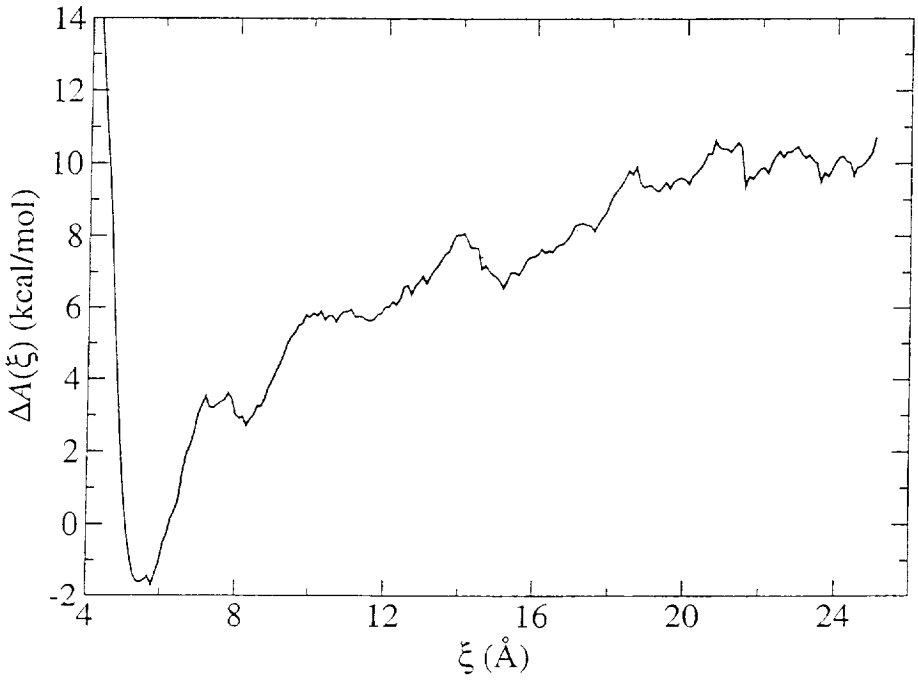

Figure 4:

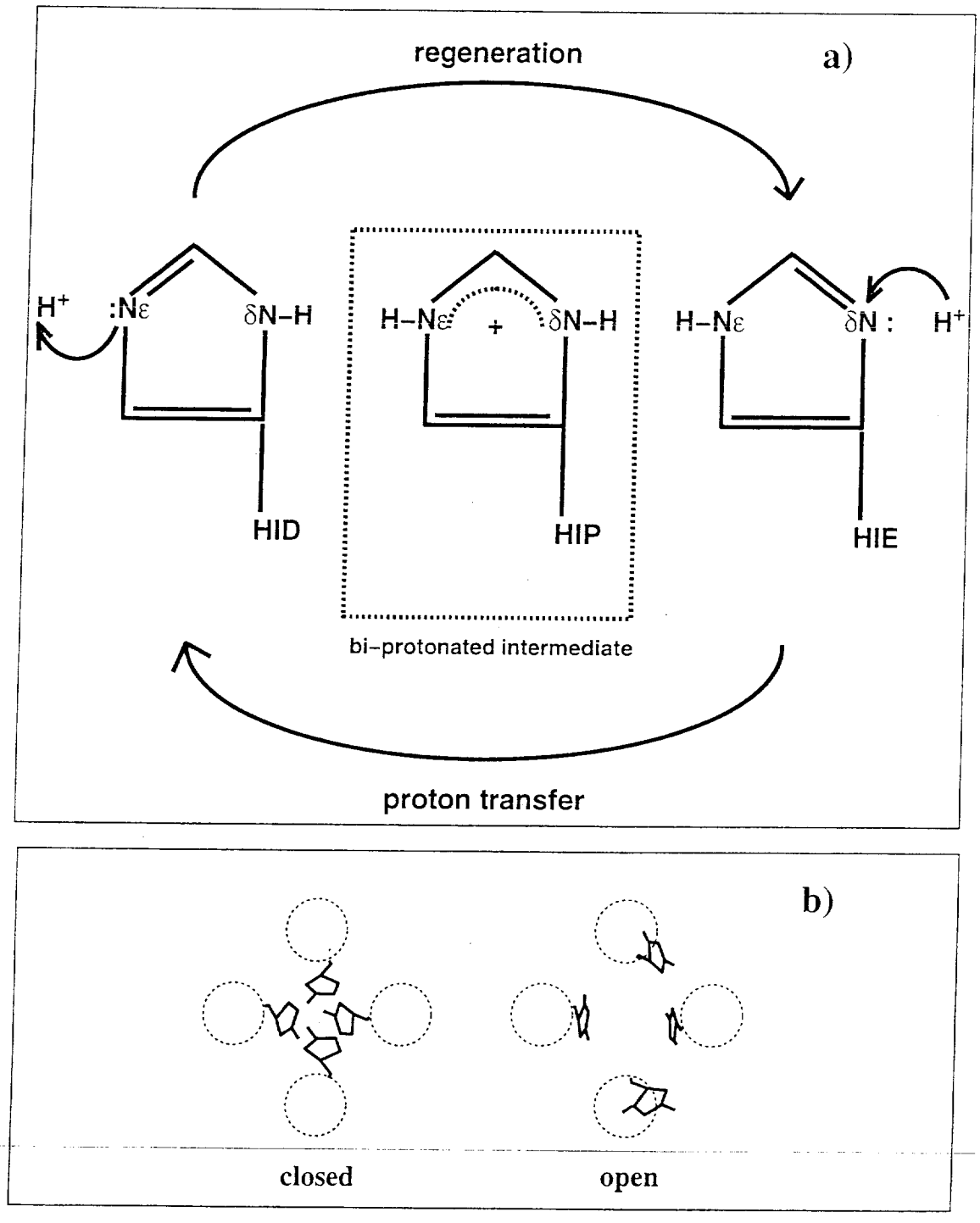
a)
regeneration
H+
:Nε
δN–H
HID
H–Nε
+
δN–H
HIP
H–Nε
δN :
H+
HIE
bi-protonated intermediate
proton transfer
b)
closed
open

Figure 5: