Matrix Formalism for Sequence-Specific Biopolymer Binding to Multicomponent Lipid Membranes*

Vladimir B. Teif,[a,b] Daniel Harries,[b] Dmitri Y. Lando,[a] and Avinoam Ben-Shaul[b]

[a]Institute of Bioorganic Chemistry, Belarus National Academy of Sciences, Minsk, Belarus

[b]Department of Physical Chemistry and The Fritz Haber Research Center, The Hebrew University, Jerusalem, Israel

Abstract

Binding of unstructured proteins to multicomponent lipid membranes is governed by a delicate interplay of electrostatic, hydrophobic, and entropic contributions. The transfer matrix formalism of statistical mechanics is a powerful way to study the effects of polymer sequence and membrane patterning using comparatively simple lattice models. We describe the general methodology and apply the method to a signal protein MARCKS binding to a three-component phosphatidyl choline / phosphatidyl serine / phosphatidylinositol 4,5-bisphosphate (PC/PS/PIP$_2$) lipid membrane. The calculated binding map shows that the protein attaches to the membrane by its myristoylated terminus and the effector domain in its middle. A bound protein sequesters about four molecules of a minor membrane lipid PIP$_2$, while the equilibrium distribution of other membrane lipids remains almost unaffected. Calculations for different mutations in the effector domain point to the importance of both lysine residues and phenyl rings inside the effector domain. Phosphorylation of three serine residues within the effector domain decreases the binding constant by three orders of magnitude and leads to protein desorption from the membrane followed by lipids redistribution.

1. Introduction

Living organisms are packed with membranes that are impermeable to water-soluble molecules acting as chemical messengers for inter- and intracellular communications. Therefore, in many cases signal transduction takes place at the membrane surface.[1,2] A key step in these processes is the binding of a protein or peptide to the lipid membrane. Lipid-peptide interactions also play a role in intracellular transport, enzyme catalysis, antimicrobial defense, and control of membrane fusion.[3] In addition to their role in vivo, membrane-active peptides are increasingly used in

pharmaceutical applications.[4] All these roles emphasize the need for a quantitative treatment of sequence-specific membrane-polymer binding. Here, we focus on unstructured amphitropic biopolymers and describe a systematic method to calculate their binding to multicomponent lipid membranes and applications of this method to signal transduction processes.

2. Biology of Membrane Binding

The cell membrane is a two-dimensional liquid composed of multiple lipid species, some of which, in addition to their structural roles, are also the precursors of second messengers. Protein binding to the membrane surface may involve both nonspecific hydrophobic and electrostatic interactions, as well as the recognition of specific lipids. Given the tabulated affinities of individual amino acids to membrane lipids,[5] one can try to predict preferred peptide conformations on a membrane. However, the computations required to solve a three-dimensional protein structure are obviously enormous. Furthermore, the computational complexity is substantially amplified by the fact that more than one molecule is usually involved in each elementary event of signal transduction. The situation is somewhat simpler in the case of unstructured peptides or proteins. Some protein segments may adopt an extended unfolded conformation upon interaction with the membrane, while the others are intrinsically unfolded in the native protein state.

The binding behavior of unstructured biopolymers is encountered, for instance, in the membrane adsorption of small signaling proteins such as the MARCKS (myristoylated alanine-rich C kinase substrate) protein and its analogues.[6–8] MARCKS acts at the inner leaflet of the plasma membrane. MARCKS binding to the membrane is enhanced through anchoring by its myristoylated terminus, which inserts into the hydrophobic membrane core. The myristoyl group is one of the fatty acyl residues, which is found as an N-terminal modification of a large number

of membrane-associated proteins. A second binding center of the MARCKS protein is the effector domain composed of 25 residues. This domain contains 13 Lys and 5 Phe residues, which interact with the membrane electrostatically and hydrophobically. The MARCKS effector domain has an increased affinity to negatively charged lipids such as monovalent PS and multivalent PIP_2. The latter lipid constitutes only about 1% of membrane lipids but is a precursor of three second messengers, inositol 1,4,5-trisphosphate [$Ins(1,4,5)P_3$], diacylglycerol (DAG), and $PtdIns(3,4,5)P_3$, and it is responsible for a wide range of membrane-related phenomena such as cytoskeletal attachment, exocytosis, endocytosis, and ion-channel activation.[7] Upon binding to the membrane, MARCKS sequesters PIP_2 lipids. MARCKS bound to the membrane is then a target for other binders, such as Ca^{2+}/calmodulin or protein kinase C (PKC). PKC phosphorylates three of four serine residues inside the effector domain of MARCKS, thus changing its affinity to the membrane and leading to its desorption. Unbinding of MARCKS releases sequestered PIP_2 molecules. Several other signal proteins bind and unbind membranes in a similar way. These mechanisms are usually referred to as the "electrostatic switch,"[6,10] "pH switch,"[9,10] etc.

3. Physics of Polymer-Membrane Binding

Protein binding to membranes bears both sequence-specific and nonspecific features. The nonspecific adsorption of polyelectrolytes at oppositely charged surfaces has been studied for over two decades using various tools of statistical thermodynamics.

Using mean-field theory, it was realized early on that a delicate interplay is established between the electrostatic tendency of a charged polymer chain to adsorb on an oppositely charged membrane and the polymer's tendency to enjoy its conformational freedom in the bulk solution and to remain unadsorbed.[11] The state of the polymer is strongly dependant on the Debye length,

l_D, which defines the distance at which coulombic interactions become screened. This screening length is defined as

$$l_D = \left(\frac{\varepsilon kT}{8\pi e^2 c}\right)^{1/2}$$

and depends on the salt concentration in the solution c, as well as the temperature T, and dielectric constant ε. As a result of the interplay between entropy and energy, when salt concentration is low—the so-called weak screening limit—the polyelectrolyte chain is highly stretched. This is because the polymer's own charge causes self-repulsion, which makes it behave much like a rod with little conformational entropy. In this case there is also little conformational loss in the adsorption process, and the polymer adsorbs easily. In contrast, in the limit $l_D \to 0$, sufficiently low temperatures are required for adsorption so as to counteract the effect of the chain's entropy and ensure that this entropy is not large enough to overcome the favorable adsorption energy.

There are further important implications to polyelectrolytes being a collection of charges connected by chemical bonds. When a polymer adsorbs on a surface it may carry with it more charge than is required for neutralizing the charged surface, leaving some of its charges facing the bulk solution. This can result in *overcharging* of a charged surface by polymers, observed as a reversal of the apparent charge of a surface in the presence of polymers. This phenomenon has been derived theoretically on the mean-field as well as using more complex theory.[13–15] A review of the current state of theory concerning polyelectrolyte interactions can be found elsewhere.[12]

More recently, an additional focus has been the rearrangement of the adsorbing surface itself as a result of interactions with the polyelectrolytes. Particularly interesting are interactions with lipid membranes that in the biological context generally behave as a two-dimensional liquid, where

lipids of various types can diffuse in the membrane plane. When polymers adsorb on mixed membranes that contain charged and neutral molecules, lipids of the opposite charge will migrate to the interaction zone to gain electrostatic interactions. In this adsorption process, counterions that were previously bound to polymer and membrane can be released into the bulk solution. This mechanism of *counterion release* is an important contribution to the adoption free energy. It is therefore important to take into account all these degrees of freedom, to include (1) the conformational entropy of the polymer chain; (2) the contributions from lipid demixing once lipids segregate around polymers so that they are effectively "bound" to them; and (3) the entropy of ions that are released in the process.[17]

Finally, at some bending energy cost, such lipid membranes can elastically deform. Clearly, in the presence of a charged polyelectrolyte, the membrane may deform so as to form better contact with the polymer. In some cases the membrane may even envelop the polymer. Conversely, the polymer may, in some cases, charge the membrane to an extent that causes it to become apparently more rigid, as fluctuations in this more charged surface become highly unfavorable.[18]

Sequence-specific binding complicates all nonspecific features mentioned above and rules out any analytical solution. A common approach to sequence-specific binding is to use lattice models. Lattice models provide a way to balance between extremes of computationally expensive all-atom simulations and approximations of simple laws of mass action.

One possibility is to look at the membrane as a fixed two-dimensional lattice of reactive centers, recognized by a one-dimensional lattice of polymer residues. Different residues may form more or less favorable contacts. Therefore, in the case of successful pattern recognition, the two-dimensional membrane pattern and the polymer sequence predefine a stable polymer configuration on the membrane. This is a sharp transition that resembles protein folding, with only minor fluctuations in the three-dimensional polymer loops, facing the bulk solution,

between several strong binding sites.[19] When there is no good match between the polymer sequence and membrane pattern, the adsorption is a smooth transition.

A different situation is encountered when the membrane residues are not fixed but are freely moving in two dimensions. This is the case for biological membranes composed of several lipid types. Now, one can describe the polymer as a one-dimensional lattice of units, which is either in three-dimensional bulk solution or in two-dimensional membrane "solution," where different lipids bind immersed polymer residues with different affinities.[20] The idea of this method is to describe polymer-membrane binding using a combination of one-dimensional models of the type widely used in molecular biology of DNA and actin[21–24] with scaling approaches of polymer physics.[19,25] Lattice models of this kind allow one to concentrate on site-specific effects abstracting from three-dimensional structure, while scaling arguments enable the introduction of simple corrections to the binding affinities arising from conformational (entropic) constraints. The latter corrections result in nonadditive binding energies of individual binding sites. For example, when an unbound polymer segment separates two binding sites, polymer looping should be taken into account. The effective binding constant of a polymer is thus determined by the binding constants of the individual segments and their positions along the sequence, as well as the membrane lipid composition.

4. The Transfer Matrix Method

4.1. General Comments

The transfer matrix methodology for calculation of sequence-specific unstructured polymer-membrane binding is based on the matrix formalism previously described as a systematic tool for the calculation of DNA-protein-drug binding in gene regulation.[23] In the context of protein-DNA

interactions, the method allows solution of many complex scenarios of cooperative assembly of proteins on the DNA molecule, which may loop and form compact structures. On the other hand, in the membrane-protein interaction, the flexible protein is the one-dimensional analogue of DNA, and the binding ligands are the (mobile) lipids embedded in the two-dimensional membrane. Assuming that the membrane is much larger than the unstructured polymer in its vicinity, we may center on the polymer and implement one-dimensional equilibrium models.[20]

In the transfer matrix formalism, the polymer is treated as a one-dimensional array of units (binding sites, residues, segments), where each unit is characterized by a matrix of statistical weights corresponding to all its possible states. Since the polymer's segments are connected, the state of a given segment depends on the states of the other segments. The system's partition function is given by successive multiplication of all transfer matrices corresponding to the polymer units. The general methodology consists of enumerating all possible states of the elementary unit, constructing the corresponding transfer matrices, and building the partition function. The partition function then allows us to calculate the maps of binding, or binding curves, and other structural and thermodynamic properties.

4.2. States Enumeration

We model the protein or peptide molecule as a linear lattice of N units labeled $n = 1, \ldots, N$. Throughout this paper we will assume that the elementary unit of the polymer is one amino acid residue interacting with membrane lipids via electrostatic or hydrophobic potentials or both. The elementary unit of the membrane is one lipid molecule.

We must list all available states for each elementary polymer unit. In the nearest-neighbor approximation, a state of a given polymer segment would depend only on its two nearest neighbors. However, we consider here a more general formalism, which allows long-range

interactions between the segments. This complication is required because the polymer may form a loop that contains many segments, and these segments "know" that they belong to the loop between the membrane-bound regions. Figure 1.1 illustrates one possible polymer-membrane configuration. We assume that the size of the lipids allows one amino acid to cover one lipid, reasonably consistent with the membrane lipids' geometry.[26] A given unit of the bound polymer may either be in contact with a lipid molecule, or reside inside a polymer loop, or belong to one of two polymer tails. Table 1.1 lists possible states for each individual unit of the polymer interacting with a membrane containing three different lipid types.
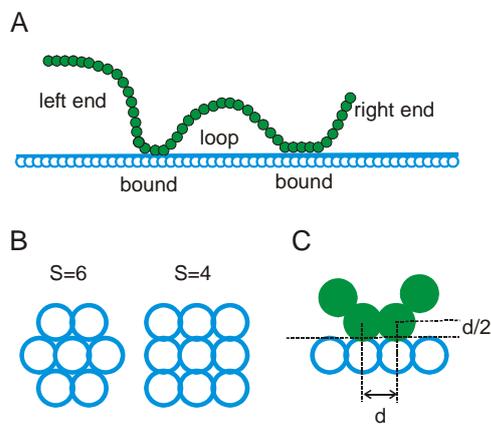
Figure 1. A) Schematic view of a polymer (green) adsorbed on a membrane (blue). B) The hexagonal and cubic lattice models. C) Geometry of polymer-membrane contacts. One polymer residue covers one membrane lipid. The lipid charges are assumed to be at the lipid-water interface; the polymer charges are at the geometrical centers of the amino acid residues. Membrane lipids and polymer residues are represented as spheres of equal diameter d.

A free polymer in three dimensions may be found in one of $C_{3D} \times (S_{3D})^N$ conformations, where N is the polymer length, C_{3D} is a constant that depends on the physical properties of the chain (e.g., flexibility, volume interactions), and S_{3D} is the number of possible orientations of a given polymer segment relative to its previous segment in a lattice model. We set the energy of a free polymer as a reference zero level, and we consider below the polymer conformations in the vicinity of the membrane when at least one polymer unit is in contact with the membrane.

4.3. Bound Polymer Segments

The statistical weights of bound polymer units are given as follows:

$$Q_n(i, j) = K_{ig}^{(n)} \times c_{0g} \times w_{gh} \times S_{2D}$$ (1.1)

Here $K_{ig}^{(n)}$ is the binding constant for the nth polymer unit of type i and a lipid of type g, $K_{ig}^{(n)} = \exp(-E_{ig}^{(n)})$, $E_{ig}^{(n)}$ is the corresponding interaction energy, and c_{0g} is the relative concentration (mole fraction) of g-type lipids in the membrane. The parameter w_{gh} is the cooperativity constant determined by the interactions between the neighboring lipids of types g and h. If g and h do not interact, $w_{gh} = 1$. S_{2D} is the statistical weight taking into account the degrees of freedom of a membrane-bound polymer residue. Since the polymer may slide along the membrane, and the lipids of the liquid membrane may move in the membrane plane, the bound polymer residue is not fixed to a point in three dimensions but is restricted to motion in two dimensions. Therefore, S_{2D} is simply the coordination number of the two-dimensional lattice model describing the membrane. In our calculations we imply hexagonal lattices with $S_{2D} = 6$ nearest neighbors in two dimensions and $S_{3D} = 12$ in three dimensions (Figure 1.1B). Because the behavior of the system is determined by the ratio S_{3D} / S_{2D} rather than by their absolute values, we may set $S_{3D} = 1$ and $S_{2D} = 0.5$.

4.4. Electrostatic Corrections to the Binding Constants

If modeling a polymer as a number of isolated noninteracting residues (beads on a string), $K_{ig}^{(n)}$ for individual beads would depend only on the type of the polymer residue i and the lipid type g, but not on the residue position n. However, in our case the residue's position should be taken into account since the entropy depends on the polymer connectivity and the energy depends on electrostatic interactions between the neighboring residues. For example, let us consider the Debye-Hückel interaction between charged polymer residues and membrane lipids. Let us

represent amino acid residues and head groups of membrane lipids as spheres with equal radii r, as shown in Figure 1.1C. Then the electrostatic energy of interaction of an isolated amino acid with an isolated lipid would be given as follows:

$$E_{ig} = z_i z_g l_B [\exp(-kr)/r], \quad K_{ig}^{(n)} = \exp(-E_{ig}) \tag{1.2}$$

where $l_B = e^2/(\varepsilon \cdot k_B T)$, the Bjerrum length, and $l_D = k^{-1}$, the Debye screening length (the typical value for the Debye length is $k^{-1} = 10$ Å and the typical distance between the centers of lipid head groups is $d = 2r = 8.66$ Å);[17] e denotes the elementary charge, z_i and z_g are the charges of the amino acid residue and the lipid, ε is the dielectric constant, k_B is the Boltzmann constant, and T is the temperature.

The first correction to eq 1.2 comes from the interaction of the lipid covered by the nth polymer unit with polymer units at positions n - 1 and n + 1. Taking these interactions in the Debye-Hückel form and using simple geometrical considerations shown in Figure 1.1C, we derive the following correction for the binding constant:

$$K_{ig}^{(n)*} = K_{ig}^{(n)} \times \sqrt[3]{K_{ig}^{(n-1)}} \times \sqrt[3]{K_{ig}^{(n+1)}} \tag{1.3}$$

Here the binding constants on the right-hand side of eq 1.3 are calculated according to eq 1.2 for isolated amino acid residues. The binding constant on the left-hand side of eq 1.3 is the corrected binding constant. The exponent 1/3 is determined by the Debye-Hückel interactions for the geometry shown in Figure 1.1C, assuming that polymer residues $n - 1$ and $n + 1$ touch the membrane. In a similar way, binding of the nth polymer residue is not only determined by the lipid underneath but also affected by the nearest lipid neighbors (six neighbors for a hexagonal lattice):

$$K_{ij}^{(n)**} = K_{ij}^{(n)*} \times \left(\sum_g c_{0g} \sqrt[3]{K_{ig}} \right)^6 \qquad (1.4)$$

4.5. Entropy of Lipid Sequestration

Let us consider a multicomponent lipid membrane. We assume that all lipid head groups are of equal size and that the monovalent and multivalent lipids interact with charged peptide units according to their charges. It is known from experiments on MARCKS binding to three-component $PC/PS/PIP_2$ membranes that PIP_2 but not PC or PS lipids are sequestered by the MARCKS(151–175) peptide.[27] PIP_2 sequestration transfers a lipid molecule from the membrane region of average concentration c_{03} to a membrane region covered by the bound peptide raising its concentration to $c_{03}*$. This changes the lipid's entropy by $\Delta S = -\ln(c_{03}*/c_{03})$.[17] Correspondingly, the binding constant decreases by a factor of $c_{03}*/c_{03}$. It is possible to find $c_{03}*$ self-consistently, changing the input parameter $c_{03}*$ until it converges to the calculated output value c_3 of the lipid bound to the polymer. For example, for MARCKS(151–175) binding to the $PC/PS/PIP_2$ membrane (89%:10%:1%) at physiological pH values, our calculation gives $c_{03}* \sim 0.16 = 4/25$, which is equivalent to about four PIP_2 molecules sequestered by the 25-residue peptide, consistent with the experimental results.[7,8,20]

4.6. Loops and Tails

The statistical weight of a loop is determined by the probability of loop formation. The probability of formation of a polymer loop of j units, starting and ending at the membrane points separated by l membrane units, is given as follows:[28]

$$P(j,l) = \frac{C}{j^{3/2}} \exp(-\frac{l^2}{2j}) \qquad (1.5)$$

Here C is the normalization constant and $j = 0, 1, \ldots, N$ is the loop length. The accurate estimation of the end-to-end loop distance is a delicate question.[28] However, for our purposes, we may simplify the loop model and treat the polymer as a random walk, and the average end-to-end distance of a polymer of length j would be simply $l = j^{1/2}$. This brings the statistical weight for the loop formation to the form $P(j) = C_{LOOP} \times j^{-\alpha}$, where C_{LOOP} is a constant dependent on loop flexibility. In particular, setting $C = 1$, we get $C_{LOOP} = \exp(-1/2) = 0.6$ for an ideal flexible chain. The loop exponent α depends on the geometry of the system: $\alpha = 1.5$ for an ideal flexible chain in three dimensions;[19] α increases if self-avoidance is taken into account and is also changed in the presence of an impenetrable membrane (see below).

Similar to end-grafted polymers, the membrane-bound polymer has one or two free ends. The weight of a free polymer end of length j is given by $(S_{3D})^j \times j^{-\beta}$, where $\beta = 0.3$ for impenetrable membranes.[25] The statistical weight of a polymer loop starting and ending on the membrane surface should also take into account the impenetrability of the membrane. Thus, the statistical weight of a first unit starting the membrane-bound polymer loop of length j is finally given as follows:

$$Q_n(i, j) = (S_{3D})^j \times C_{LOOP} \times j^{-(\alpha+\beta)} \tag{1.6}$$

4.7. Transfer Matrix Construction

The element $Q_n(i, j)$ of the transfer matrix Q_n expresses the statistical weight corresponding to the nth polymer unit in state i, followed by the next unit in state j. Note that only specific combinations of states i and j are allowed. Allowed states are characterized by statistical weights depending on the concentrations and energetic parameters, as detailed above. Forbidden states

are characterized by zero statistical weights. We consider three types of membrane lipids as in Table 1.1. The transfer matrices $Q_{ij}^{(n)}$ are then constructed according to the following algorithms:

(1) Bound unit followed by bound unit ($i = 1, \ldots, 3, j = 1, \ldots, 3$):

$$Q_{ij}^{(n)} = K_{ij}^{(n)} \times c_{0i} \times S_{2D} \times w_{ij}$$

(2) Bound unit followed by a loop ($i = 1, \ldots, 3, j = 4, \ldots, N + 2$):

$$Q_{ij}^{(n)} = K_{ij}^{(n)} \times c_{0i} \times S_{2D} \times C_{LOOP} \times (j - 3)^{-\alpha-\beta}, \text{ if } j - 4 < N - (n + 1)$$

(3) Bound unit followed by the free right polymer end ($i = 1, \ldots, 3, j = N + 4$):

$$Q_{ij}^{(n)} = K_{ij}^{(n)} \times c_{0i} \times S_{2D} \times (N - n)^{-\beta}, \text{ if } n < N$$

(4) Free left polymer end followed by bound unit ($i = 4, j = 1, \ldots, 3$):

$$Q_{ij}^{(n)} = 1, \text{ if } 1 < n < N$$

(5) Loop continues ($i = 5, \ldots, N + 2, j = i - 1$):

$$Q_{ij}^{(n)} = 1, \text{ if } n > 1 \text{ and } i - 4 < N - n$$

(6) Left end continues ($i = N + 3, j = 1, \ldots, 3$):

$$Q_{ij}^{(n)} = 1, \text{ if } n < N - 1$$

(7) Right end continues ($i = N + 4, j = N + 4$):

$$Q_{ij}^{(n)} = 1, \text{ if } n > 1$$

The remaining matrix elements are zeroes. Nonzero elements are shaded in Figure 1.2.

n+1 / n	1	2	3	4	5	...	N+2	N+3	N+4
1						...			
2						...			
3						...			
4						...			
5						...			
...	...	...	...	...	...	...	...	...	...
N+2						...			
N+3						...			
N+4						...			

Figure 2. The transfer matrix for a residue of a polymer of length N, binding to a three-component lipid membrane. The non-zero matrix elements are shaded.

4.8. Calculating Binding Probabilities

The matrices constructed above correspond to regular polymer residues far from the ends and from other obstacles. Close to the polymer ends, transfer matrices change according to boundary conditions. For example, the loop cannot propagate beyond the polymer ends—thus the loop of length j cannot start within the last $j - 1$ polymer units. Our boundary conditions also imply that the first transfer matrix is preceded by the vector $(1, 1, …, 1)$, and the last matrix is followed by the vector $(1, 1, …, 1)^T$. This is required to get a scalar value of the partition function as a final result of the matrix multiplication. The partition function Z and its derivatives are calculated using the recursive multiplication of all transfer matrices according to the polymer sequence:[22]

$$Z = A_N \times \begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix}, \quad A_i = A_{i-1} \times Q_n, \quad A_0 = \begin{pmatrix} 1 & 1 & ... & 1 \end{pmatrix} \tag{1.7}$$

$$\frac{\partial Z}{\partial X} = \frac{\partial A_N}{\partial X} \times \begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix}, \quad \frac{\partial A_n}{\partial X} = \frac{\partial A_{n-1}}{\partial X} \times Q_n + A_{n-1} \times \frac{\partial Q_n}{\partial X}, \quad A_0 = \begin{pmatrix} 1 & 1 & ... & 1 \end{pmatrix} \tag{1.8}$$

The probability c_{ng} that the nth polymer segment is bound to a lipid molecule of type g is given by the following equation:[23]

$$c_{ng} = \frac{\partial Z}{\partial K_{ig}^{(n)}} \times \frac{K_{ig}^{(n)}}{Z} \qquad (1.9)$$

The whole set of c_{ng} values determines the map of lipid binding to the polymer. The probability c_n that the nth polymer unit is bound to the membrane (to any membrane lipid) is $c_n = \sum_g c_{ng}$, and the number of sequestered g-type lipids is $c_g = \sum_n c_{ng}$.

The membrane-polymer binding constant (also known as the partition coefficient) is given by the ratio between the partition functions of free and bound polymer conformations: $K = Z_{bound} / Z_{free}$, where Z_{bound} is the partition function calculated as described above, and Z_{free} is the partition function calculated for the polymer of the same length, setting all the energies of polymer-lipid interaction equal to zero.

5. Constructing the Matrix Model for the MARCKS Protein

Binding of signal proteins like MARCKS and related peptides to mixed lipid membranes was studied previously by solving the Poisson-Boltzmann equation[29,30,36] and using Monte Carlo simulations.[17] In our recent work, we have applied the transfer matrix method to peptides corresponding to the MARCKS effector domain and a series of related peptides.[20] Here we provide matrix calculations for the whole sequence of the MARCKS protein.

5.1. Types of Elementary Units

The human MARCKS protein consists of 331 amino acids, as shown in Figure 1.3. It is possible to distinguish six types of residues:

- $i = 1$—neutral nonaromatic (Leu, Ser, Gly, Ala)

- $i = 2$—basic (Lys and Arg, charge = +1)

- $i = 3$—aromatic (Phe)

- $i = 4$—phosphorylated Ser (charge = –2)

- $i = 5$—acidic (Glu, charge = –1)

- $i = 6$—myristate anchor (neutral, hydrophobic)

We focus on cell membranes composed of three types of lipids:

- $g = 1$—neutral PC

- $g = 2$—monovalent PS; charge = –1)

- $g = 3$—multivalent PIP_2 (charge varies from –3 to –5 in different experiments; we take charge = –4, corresponding to the physiological pH range)[8]

GAQFSKTAAKGEAAAERPGEAAVASSPSKANGQENGH
VKVNGDASPAAAESGAKEELQANGSAPAADKEEPAAA
GSGAASPSAAEKGEPAAAAAPEAGASPVEKEAPAEGE   Figure 3. Sequence of the human MARCKS protein
AAEPGSPTAAEGEAASAASSTSSPKAEDGATPSPSNE
TPKKKKKRFSFKKSFKLSGFSFKKNKKEAGEGGEAEA
PAAEGGKDEAAGGAAAAAAEAGAASGEQAAAPGEEAA
AGEEGAAGGDPQEAKPQEAAVAPEKPPASDETKAAEE
PSKVEEKKAEEAGASAAACEAPSAAGPGAPPEQEAAP
AEEPAAAAASSACAAPSQEAQPECSPEAPPAEAAE

5.2. Choosing Parameters

We set to zero the interaction energy of neutral nonaromatic polymer units with the membrane. Upon interaction with the membrane, aromatic residues are buried in the hydrophobic core of the lipid bilayer, while charged units remain primarily in the water phase. Phe residues bind membranes with energy ranging from 0.2 kcal/mol[31] to 1.3 kcal/mol,[5] depending on experimental conditions. Fitting of experimental adsorption isotherms of the MARCKS effector domain results in 0.8 kcal/mol.[20]

The energy of myristate binding to a lipid membrane is well determined experimentally.[32] Experiments give energy of 0.8 kcal/mol per each carbon of the 14-carbon myristate, and finally the binding constant for the myristoyl "anchor" is $10^4\,M^{-1}$. It is interesting to note that the 0.8 kcal/mol value of each carbon of the myristate is similar to the energy of each Phe residue mentioned above. This similarity is explained by hydrophobic mechanisms, which dominate in both cases.

The strength of electrostatic interactions depends on our choice of the effective dielectric constant near the membrane (ranging between $\varepsilon = 78$ in water and $\varepsilon = 2$ inside the membrane). Our fitting of experimental data on the MARCKS(151–175) peptide gives an effective dielectric constant of $\varepsilon = 55$.[20] We use these values of electrostatic and hydrophobic interactions in calculations without further fitting. The choice of the reference experimental data set does not affect the relative changes in the binding affinities arising because of the changes in the membrane composition and polymer sequence. We assume that the distance between the centers of lipid head groups is $d = 8.66$ Å, which is also used to model the bond length between the neighboring polymer segments.[17] The charged lipids and Lys and Arg polymer residues interact through the Debye-Hückel potential corrected according to eqs 1.2 through 1.4.

6. Calculating MARCKS-Membrane Binding

Figure 1.4 shows the binding map calculated for the human MARCKS and a three-component PC/PS/PIP$_2$ (89%:10%:1%) lipid membrane. The binding map plots the probability for each polymer residue to be bound to each type of membrane lipid. Protein residues are numbered from left to right according to Figure 1.3. The calculated binding map in Figure 1.4 shows the

following features: (1) the polymer is preferentially bound to the membrane at the myristoylated end and several nearby residues; (2) the farther the polymer region is situated from the myristate anchor, the less the probability of binding; (3) a second strong binding center is at the basic effector domain in the middle of the protein; (4) the effector domain contains three strong sites for PIP_2 sequestration; and (5) an additional PIP_2 molecule may be bound outside the effector domain, delocalized among the small green peaks in Figure 1.4.
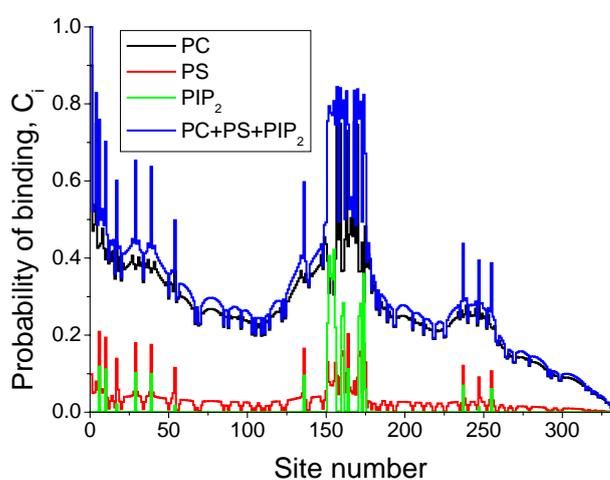
Figure 4. The map of binding of the MARCKS protein to the $PC/PS/PIP_2$ (89% : 10% : 1%) membrane. Each protein residue is characterized by a probability to be bound to a PC lipid (black line), PS (red), PIP_2 (green) and any of these lipids (blue).

Figure 1.5 shows our calculation of the protein-membrane binding constant. Here we again take a three-component $PC/PS/PIP_2$ membrane containing 1% PIP_2 but change the percentage of the monovalent lipid PS. We have performed calculations for different mutations of the MARCKS protein.
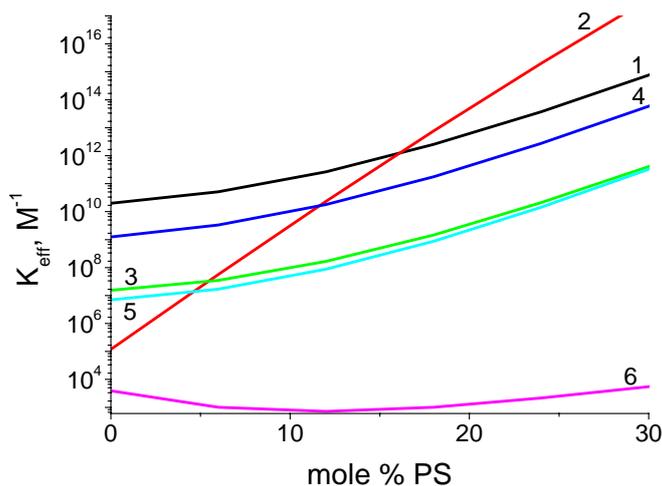
Figure 5. The effective binding constant calculated for different mutations of the MARCKS protein to the PC/PS/PIP2 (89/10/1) membrane. 1 – the human MARCKS sequence; 2 – Glu residues are substituted by Ala; 3 – the myristate moiety has been deleted; 4 – one Ser residue inside the effector domain is phosphorylated; 5 – three Ser residues are phosphorylated; 6 – the 25-residue effector domain is removed.

Curve 1 in Figure 1.5 corresponds to the native human MARCKS protein. The curve starts at a high value determined by the hydrophobic anchoring of the myristoyl moiety and further increases upon loading negatively charged PS lipids. Our previous calculations for the 25-residue peptide corresponding to the MARCKS effector domain showed a linear increase of K_{eff} with increasing PS fraction.[20] That the increase is not linear for the native MARCKS protein indicates that there are two opposite electrostatic contributions. Negative lipids attract basic Lys residues of the effector domain, but they repel acidic Glu residues randomly scattered through other parts of the protein. We have checked this by substituting all acidic Glu residues by neutral Ala. Curve 2 in Figure 1.5 shows the calculation for the MARCKS protein with the Glu→Ala substitution. Now there is no repulsive effect, and the binding constant indeed increases linearly with increased content of negative lipids in the membrane.

Curve 3 corresponds to the MARCKS mutation when the myristate has been deleted from its end. The trend of this curve is similar to the native MARCKS protein but is shifted down by about three orders of magnitude. The shift value is comparable to the value of the myristate-

membrane binding constant but is not exactly equal to it, reflecting the fact that the binding constants of different polymer parts are nonadditive.

Curve 4 in Figure 1.5 shows calculations for the MARCKS protein with one phosphorylated Ser residue. This covalent modification adds a charge of –2, which is not a large change for a protein whose total charge is around –30. However, as we can see from the figure, this shifts the effective binding constant by an order of magnitude. This shift is because a single phosphorylation changes the local charge distribution inside the highly charged basic effector domain. A decrease in the binding constant of three orders of magnitude is induced when three Ser residues are phosphorylated (curve 5).

Curve 5 corresponds to the situation of three out of four Ser residues inside the MARCKS effector domain being phosphorylated. This situation has direct biological implications, since PKC usually triply phosphorylates membrane-bound MARCKS. It is this chemical modification that leads to MARCKS dissociation from the membrane followed by the release of sequestered PIP$_2$ lipids. Our calculations predict that phosphorylation of three Ser residues decreases MARCKS binding constant by three orders of magnitude, possibly explaining its desorption from the membrane. Monte Carlo simulations provide a similar estimate.[37] Interestingly, the effect of phosphorylation (curve 5) is similar to the effect of deletion on the myristate moiety (curve 3).

Curve 6 corresponds to the most dramatic mutation of the MARCKS sequence, when 25 residues of the effector domain are removed. This modification changes the MARCKS binding constant by six orders of magnitude, which practically guarantees its desorption from the membrane. The electrostatic attraction now turns to repulsion. Only the myristoyl anchor can attach the protein, which by itself is not enough to compensate for the entropy decrease due to protein confinement.

The deletion of the effector domain also makes the protein less sensitive to membrane composition. Since now there is almost no electrostatic contribution to the binding energy, increasing the percentage of charged membrane lipids causes negligible changes in the binding constant. Although the deletion of the effector domain by itself is not a biological event, it may help us to estimate nonelectrostatic results of PKC-induced MARCKS modifications. Indeed, in addition to the changes in electrostatics, phosphorylation may cause structural changes in the effector domain that are mediated, for example, by salt bridges between phosphorylated serine and lysine,[33] which would prohibit its binding to the membrane. We see, therefore, that Figure 1.5 adequately explains the action of PKC on MARCKS-membrane binding and unbinding behavior.

7. Potential Applications to More Complex Systems

We have presented a general method for calculating sequence-specific binding of flexible peptides and unstructured proteins to mixed lipid membranes. The calculations for the MARCKS protein allowed us to determine relative changes in the binding constant arising because of the changes in the peptide sequence and membrane composition, in agreement with experiments. The matrix method may be easily extended to consider the binding of second-layer molecules to the proteins already bound to the membrane.[23] We therefore hope that it is not only applicable to study signal transduction through single-protein binding to a membrane but also may be extended in future to include multiprotein assemblies on the membrane. This could help in studies of, for example, the membrane-cytoskeleton attachment[34] and its regulation by binding of small ligands such as adenosine 5′-triphosphate (ATP) and Ca^{2+}. Multilayer matrix models may be also applicable to lipid-templated amyloid-type protein fibril formation.[35]

References

1. Eyster KM. **The membrane and lipids as integral participants in signal transduction: Lipid signal transduction for the non-lipid biochemist.** *Adv Physiol Educ* 2007, **31**:5–16.

2. Cho W and Stahelin RV. **Membrane-protein interactions in cell signaling and membrane trafficking.** *Annu Rev Biophys Biomol Struct* 2005, **34**:119–151.

3. Dowhan W, Mileykovskaya E, and Bogdanov M. **Diversity and versatility of lipid-protein interactions revealed by molecular genetic approaches.** *Biochim Biophys Acta* 2004, **1666**:19–39.

4. Henriques ST, Melo MN, and Castanho MA. **Cell-penetrating peptides and antimicrobial peptides: How different are they?** *Biochem J* 2006, **399**:1–7.

5. Wimley WC and White SH. **Experimentally determined hydrophobicity scale for proteins at membrane interfaces.** *Nat Struct Biol* 1996, **3**:842–848.

6. McLaughlin S and Aderem A. **The myristoyl-electrostatic switch: A modulator of reversible protein-membrane interactions.** *Trends Biochem Sci* 1995, **20**:272–276.

7. McLaughlin S and Murray D. **Plasma membrane phosphoinositide organization by protein electrostatics.** *Nature* 2005, **438**:605–611.

8. McLaughlin S, Wang J, Gambhir A, and Murray D. **PIP(2) and proteins: Interactions, organization, and information flow.** *Annu Rev Biophys Biomol Struct* 2002, **31**:151–175.

9. Hanakam F, Gerisch G, Lotz S, Alt T, and Seelig A. **Binding of hisactophilin I and II to lipid membranes is controlled by a pH-dependent myristoyl-histidine switch.** *Biochemistry* 1996**, 35**:11036–11044.

10. Johnson JE and Cornell RB. **Amphitropic proteins: Regulation by reversible membrane interactions (review).** *Mol Membr Biol* 1999, **16**:217–235.

11. Muthukumar M. **Adsorption of a polyelectrolyte chain to a charged surface.** *J Chem Phys* 1987, **86**:7230–7235.

12. Netz RR and Andelman D. **Neutral and charged polymers at interfaces.** *Phys Rep* 2003, **380**:1–95.

13. Joanny JF. **Polyelectrolyte adsorption and charge inversion.** *Eur Phys J B* 1999, **9**:117–122.

14. May S, Harries D, and Ben-Shaul A. **Lipid demixing and protein-protein interactions in the adsorption of charged proteins on mixed membranes.** *Biophys J* 2000, **79**:1747–1760.

15. Grosberg A-Yu, Nguyen TT, and Shklovskii BI. **Low temperature physics at room temperature in water: Charge inversion in chemical and biological systems**, *Rev Mod Phys* 2002, **74**:329–345.

16. Dobrynin AV, Deshkovski A, and Rubinstein M. **Adsorption of polyelectrolytes at an oppositely charged surface.** *Macromolecules* 2001, **34**:3421–3436.

17. Tzlil S and Ben-Shaul A. **Flexible charged macromolecules on mixed fluid lipid membranes: theory and Monte Carlo simulations.** *Biophys J* 2005, **89**:2972–2987.

18. Shafir A and Andelman D. **Bending moduli of charged membranes immersed in polyelectrolyte solutions**. *Soft Matter* 2006, **3**:644–650.

19. Chakraborty AK. **Disordered heteropolymers: models for biomimetic polymers and polymers with frustrating quenched disorder.** *Phys Rep* 2001, **342**:1–61.

20. Teif VB, Harries D, Lando DY, and Ben-Shaul A. **Matrix formalism for site-specific binding of unstructured proteins to multicomponent lipid membranes.** *J Pept Sci*, 2008, 4:368–373.

21. Crothers DM. **Calculation of binding isotherms for heterogenous polymers.** *Biopolymers* 1968, **6**:575–584.

22. Zasedatelev AS, Gurskii GV, and Vol'kenshtein MV. **Theory of one-dimensional adsorption. I. Adsorption of small molecules on a homopolymer.** *Mol Biol* 1971, **5**:194–198.

23. Teif VB. **General transfer matrix formalism to calculate DNA-protein-drug binding in gene regulation: application to O_R operator of phage λ.** *Nucleic Acids Res* 2007, **35**:e80.

24. von Hippel PH. **From "simple" DNA-protein interactions to the macromolecular machines of gene expression.** *Annu Rev Biophys Biomol Struct* 2007, **36**:79–105.

25. Slutsky M, Zandi R, Kantor Y, and Kardar M. **Apex exponents for polymer-probe interactions.** *Phys Rev Lett* 2005, **94**:198303. {AU: page numbers correct?} {ED: this journal has article numbers instead of page numbers. 198303 is the article number}

26. Huang C and Mason JT. **Geometric packing constraints in egg phosphatidylcholine vesicles.** *Proc Natl Acad Sci USA* 1978, **75**:308–310.

27. Golebiewska U, Gambhir A, Hangyas-Mihalyne G, Zaitseva I, Radler J, and McLaughlin S. **Membrane-bound basic peptides sequester multivalent (PIP₂), but not monovalent (PS), acidic lipids.** *Biophys J* 2006, **91**:588–599.

28. Chakraborty AK and Bratko D. **A simple theory and Monte Carlo simulations for recognition between random heteropolymers and disordered surfaces.** *J Chem Phys* 1998, 108:1676–1682.

29. Ben-Tal N, Honig B, Miller C, and McLaughlin S. **Electrostatic binding of proteins to membranes. Theoretical predictions and experimental results with charybdotoxin and phospholipid vesicles.** *Biophys J* 1997, **73**:1717–1727.

30. Haleva E, Ben-Tal N, and Diamant H. **Increased concentration of polyvalent phospholipids in the adsorption domain of a charged protein.** *Biophys J* 2004, **86**:2165–2178.

31. Victor K, Jacob J, and Cafiso DS. **Interactions controlling the membrane binding of basic protein domains: Phenylalanine and the attachment of the myristoylated alanine-rich C-kinase substrate protein to interfaces.** *Biochemistry* 1999, **38**:12527–12536.

32. Peitzsch RM and McLaughlin S. **Binding of acylated peptides and fatty acids to phospholipid vesicles: pertinence to myristoylated proteins.** *Biochemistry* 1993, **32**:10436–10443.

33. Errington N and Doig AJ. **A phosphoserine-lysine salt bridge within an α-helical peptide, the strongest α-helix side-chain interaction measured to date.** *Biochemistry* 2005, **44**:7553–7558.

34. Sheetz MP, Sable JE, and Dobereiner HG. **Continuous membrane-cytoskeleton adhesion requires continuous accommodation to lipid and cytoskeleton dynamics.** *Annu Rev Biophys Biomol Struct* 2006, **35**:417–434.

35. Murphy RM. **Kinetics of amyloid formation and membrane interaction with amyloidogenic proteins.** *Biochim Biophys Acta* 2007, **1768**:1923–1934.

36. Khelashvili G, Weinstein H, and Harries D. **Protein diffusion on charged membranes: A dynamic mean-field model describes time evolution and lipid reorganization.** *Biophys J* 2008, **94**:2580–2597.

37. Tzlil S, Murray D, and Ben-Shaul A. The "electrostatic-switch" mechanism: Monte Carlo study of MARCKS-membrane interaction. *Biophys J* 2008, 95:1745–1757.