Effect of Protein Aggregation in the Aqueous Phase on the Binding of Membrane Proteins to Membranes

Robert Doebler,* Nilay Başaran,# Harold Goldston,* and Peter W. Holloway*
*Department of Biochemistry and Molecular Genetics, Health Sciences Center, and the Biophysics Program, University of Virginia, Charlottesville, Virginia 22908 USA, and #Department of Biology, Middle East Technical University, Ankara 06531, Turkey

ABSTRACT Analysis of the binding of hydrophobic peptides or proteins to membranes generally assumes that the solute is monomeric in both the aqueous phase and the membrane. Simulations were performed to examine the effect of solute self-association in the aqueous phase on the binding of monomeric solute to lipid vesicles. Aggregation lowered the initial concentration of monomeric solute, which was then maintained at a relatively constant value at the expense of the aggregated solute, as the lipid concentration was increased. The resultant binding isotherm has a more linear initial portion rather than the classic hyperbolic shape. Although this shape is diagnostic of solute self-association in the aqueous phase, various combinations of values for the membrane partition coefficient and the solute self-association constant will generate similar isotherms. Data for cytochrome b_5 were analyzed and, when the self-association constant was estimated by gel filtration, a unique value for the membrane partition coefficient was obtained. Thus, to obtain a true partition coefficient the state of the solute in the aqueous phase must be known. If the concentration of the monomeric solute species in the aqueous phase can be independently determined, then, even with heterogeneous aggregates, the true partition coefficient can be obtained.

INTRODUCTION

It is generally accepted that the unique localization of membrane proteins in cell membranes is due in most cases to distributed segments of hydrophobicity along the polypeptide chain. Many efforts have been made to quantitate the hydrophobicity of a protein, a peptide, or even a single amino acid, by measuring the partitioning of it between water and lipid bilayers or between water and an organic phase. One ultimate aim of such studies is to place the individual amino acids on a "hydrophobicity scale." If such a scale could be accurately constructed it would be possible to reliably predict which segments in a protein sequence would be membrane-inserted. Many such scales have been generated and have been summarized extensively (Eisenberg, 1984; Wimley and White, 1996).

Several methods have been used to compare the hydrophobicities of the amino acid side chains, and the most complete experimental study is that of Wimley and White (1996), which builds upon earlier studies (Wimley and White, 1992, 1993; Jacobs and White, 1989). In their recent report they measured the mole-fraction partition coefficients between buffer and lipid vesicles of a series of "host-guest" pentapeptides, in which the central residue in the pentapeptide was each of the 20 amino acids taken in turn. As these peptides were shown not to penetrate deeply into the lipid bilayer, these studies determine the free energies of transfer of each amino acid side chain between water and the bilayer interface. Other studies have investigated the interaction of larger peptides, which penetrate deeper into the membrane, and the free energies of transfer of various mutated hydrophobic signal sequences have also been determined (Jones and Gierasch, 1994).

This report examines the effect of self-association in the aqueous phase on membrane partitioning. As far as we are aware this has not been considered previously, although a solute that has enough hydrophobic character to associate with a membrane would be expected to self-associate in the aqueous phase. Indeed, Wimley and White (1996) noted that one of the peptides they had synthesized, $AcWL_6$, appeared to aggregate in the aqueous phase and hence the authors did not make use of the ΔG of transfer calculated for this peptide in any of the computations performed in that publication. The only other case in which the effect of self-association of a peptide solute upon partitioning has been noted was in the studies of the interaction of the peptide alamethicin with lipid vesicles by Schwarz and co-workers (Rizzo et al., 1987). They noted that the binding of alamethicin to vesicles could only be simulated by postulating the self-association of the peptide in the bilayer but not by self-association of the peptide in the aqueous phase. Although this self-association in the membrane is to be expected because of the pore-forming properties of alamethicin, it has also been shown to occur with certain transmembrane helices derived from intrinsic membrane proteins (Bormann and Engelman, 1992).

In order to examine the effect of self-association in the aqueous phase on membrane partitioning of hydrophobic peptides we have simulated binding curves, by an iterative numerical technique, for a solute that undergoes self-association in the aqueous phase. The effect of the number of

Received for publication 27 April 1998 and in final form 27 October 1998.

Address reprint requests to Peter W. Holloway, Department of Biochemistry and Molecular Genetics, Health Sciences Center, University of Virginia, Charlottesville, VA 22908. Tel.: 804-924-2509; Fax: 804-924-5069; E-mail: pwh1j@virginia.edu.

Abbreviations used: b_5, native rabbit endoplasmic reticulum cytochrome b_5 isolated from *Escherichia coli* by detergent extraction; POPC, 1-palmitoyl-2-oleoylphosphatidylcholine; SUV, small unilamellar vesicles.

0006-3495/99/02/928/09 $2.00

molecules in the aggregate and the strength of the self-association were examined. Because the amount of solute bound to the vesicle is related to both the membrane partition coefficient and, reciprocally, to the strength of solute self-association in the aqueous phase, quite similar binding curves can be generated by different combinations of these two parameters. This suggests that the evaluation of an experimental partition coefficient for a self-associating solute is difficult unless the dissociation constant for the aggregate can be obtained independently. As a test of these conclusions, previously published binding data on cytochrome b_5 (b_5) (Tretyachenko-Ladokhina et al., 1993) were reanalyzed with the additional knowledge of the dissociation constant of the aqueous phase aggregate.

Cytochrome b_5 has been studied extensively as a model membrane protein since it has the advantage of being water-soluble in the absence of detergents. In aqueous solution the protein has been shown to exist predominantly as an octamer (Spatz and Strittmatter, 1971; Calabro et al., 1976), although the binding to lipid vesicles has been shown to proceed through the monomeric form (Leto and Holloway, 1979; Krishnamachary et al., 1994). It is also known from previous studies that b_5 does not aggregate in the bilayer plane when bound to vesicles (Freire et al., 1983). We are currently interested in the effect of amino acid substitutions in the membrane-binding domain and alterations in the composition of the lipid bilayer on the interaction of the protein with vesicles (Başaran et al., 1996). If the protein self-associates in the aqueous phase, then this will complicate the normal analysis of binding curves that are based on a two-state model. We show here that the fitting of an experimental b_5 binding curve without any constraints does not produce reasonable results. However, when the value of the dissociation constant for the aggregate evaluated by small-zone gel filtration is used, reasonable estimates of both the partition coefficient and the aggregation number of the aggregate are obtained. This analysis is applicable to other measurements of partition coefficients and, provided the concentration of free monomeric solute in the aqueous phase can be determined, the true partition coefficient can be obtained even when the solute is present as random aggregates in the aqueous phase.

MATERIALS AND METHODS

The native rabbit endoplasmic reticulum form of cytochrome b_5 was isolated from *Escherichia coli* as described previously (Ladokhin et al., 1991). The buffer used in the gel filtration was 10 mM HEPES-0.1 mM EDTA (pH 7.3 at 25°C), which was the same buffer used previously to study the binding of b_5 to SUV made from POPC (Tretyachenko-Ladokhina et al., 1993). The conductivity of this buffer as measured with an Omega CDB-70 conductivity meter (Omega Engineering, Stamford, CT) was 504 μS. The measurement of conductivity was a simple procedure for standardizing the ionic strength of various buffers. Samples of 10, 5, 2.5, 1.9, or 1.25 nmol of b_5 in 40 μl HEPES buffer were mixed with 10 μl 5M KCl, to ensure all the protein was in the aggregated form (Calabro et al., 1976), and were then applied to a column (1 × 22 cm) of Sephadex G200 Superfine at 25°C. The flow rate was 2.4 ml/h and the effluent from the column was monitored at 412 nm with a flow cell in a spectrophotometer, the output from which was interfaced to a computer via an interface board (Omega Engineering, Stamford, CT). Data points were taken every 20 s and written in a computer file. The concentration of protein eluting from the column was estimated using an $\epsilon = 117{,}000\ M^{-1}\ cm^{-1}$.

RESULTS AND DISCUSSION

Simulations of protein binding to vesicles

Simulations of the binding of a protein to lipid vesicles were performed using a partition model. The mole fraction partition coefficient, K_X, is defined as:

$$K_X = ([S]_L/[L])/([S]_W/[W]) \quad (1)$$

where $[S]_L$ and $[S]_W$ are the molar concentrations of solute in the vesicle fraction and aqueous fraction (bound and free), respectively, and [L] and [W] are the molar concentrations of lipid and water, respectively (Wimley and White, 1996).

If the solute undergoes self-association into an aggregated form, an N-mer, in the aqueous phase:

$$NS \rightleftharpoons S_N$$

In practice, the concentration of S_N would usually be determined by measuring the concentration of the molecules of S which are in the aggregated form. Hence the actual concentration of S_N will be the measured concentration of the solute S which is in the aggregated form, divided by N ($[S_N]_W/N$). The dissociation constant of this aggregation process can then be written as:

$$K_P = N \cdot ([S]_W)^N/[S_N]_W \quad (2)$$

where K_P is the dissociation constant for the conversion of N-mer to monomer and $[S_N]_W$ is the concentration of the solute molecules which are in the N-mer form in the aqueous phase. The existence of this equilibrium will influence the binding of the solute to lipid vesicles. For any lipid or solute concentration both the solute-lipid interaction equilibrium (Eq. 1) and the protein dissociation equilibrium (Eq. 2) must simultaneously be obeyed. In particular, the concentration of $[S]_W$ available for binding to the vesicles will depend upon the equilibrium expressed in Eq. 2.

The effect of the equilibrium in Eq. 2 upon the binding of solute to vesicles, modeled by Eq. 1, was simulated by a program written in BASIC. In this program, trial values of K_P, N, and total solute concentration were used to calculate the concentrations of free monomer and N-mer in an iterative manner using Eq. 2. This value of monomer concentration was then used with Eq. 1 to calculate the concentration of bound solute. This results in a perturbation of the equilibrium shown in Eq. 2, which then requires a recalculation of monomer and N-mer concentrations. At each lipid concentration, the values found for free monomer tended to converge in the program after ~20 iterations of Eq. 2 and the values of bound protein after 300 iterations of Eq. 1. More importantly, at this point the values of K_P and K_X calculated from the derived concentrations of the three

solute species were equal to the original values input for K_P and K_X. As the number of iterations required depended upon the lipid concentration, the number of iterations for Eq. 2 were set at 100 and the number of iterations of Eq. 1 at 1000, to ensure the correct convergence had occurred at each lipid concentration. At this point the program incremented the lipid concentration by a predetermined amount and the process was repeated.

For the initial simulations, the protein was assumed to undergo cooperative aggregation to an 8-mer. In all cases the partition coefficient (K_X) was set at 10^7 and K_P was set at infinity, with no aggregation occurring, or at 3.29×10^{-52} M^7 or 2.09×10^{-55} M^7. The choice of N came from earlier studies with b_5 where there was evidence that the protein formed an 8-mer (Calabro et al., 1976) and K_X came from preliminary fitting of b_5 binding data (see below). The values of K_P were obtained from Eq. 2 by setting the concentrations of monomeric protein, in the presence of 1 μM total protein, to realistic values of 0.05 μM and 0.02 μM (see b_5 data later). As shown in Fig. 1, when K_P = infinity (*closed circles*) a normal binding isotherm is obtained. As K_P is assigned smaller values (*closed squares* and *closed triangles*), the simulated binding curves at low lipid concentrations become more linear and the protein appears to have a lower affinity for the lipid.

The validity of the simulations can be seen from the following model calculations. The simulated data, from which the isotherm represented by the closed triangles was generated, gave the following values of the three protein "species" when 101 μM lipid was present: protein molecules present as 8-mer in the aqueous phase ($[S_N]_W$), 0.63680 μM; protein molecules present as monomer in the aqueous phase ($[S]_W$), 0.01895 μM; and protein molecules bound to lipid ($[S]_L$), 0.34425 μM. When these values are inserted into Eq. 1,

$$\begin{aligned} K_X &= ([S]_L/[L])/([S]_W/[W]) \\ &= (0.34425 \times 10^{-6}) \cdot (55.6)/(101 \times 10^{-6}) \\ &\quad \cdot (0.01895 \times 10^{-6}) \\ &= 0.9995 \times 10^7 \end{aligned}$$

compared to the input value of 10^7 for K_X. When the values are inserted into Eq. 2,

$$\begin{aligned} K_P &= N \cdot ([S]_W)^N/[S_N]_W \\ &= (8) \cdot (0.01895 \times 10^{-6})^8/(0.63680 \times 10^{-6}) \\ &= 2.089 \times 10^{-55}\, M^7, \end{aligned}$$

compared to the input value for K_P of 2.09×10^{-55} M^7. The protein concentrations in this simulation are here presented to unrealistic accuracy because any rounding of the values would have a large effect upon the ultimate calculated value of K_P, and so mask the precision of the calculations.

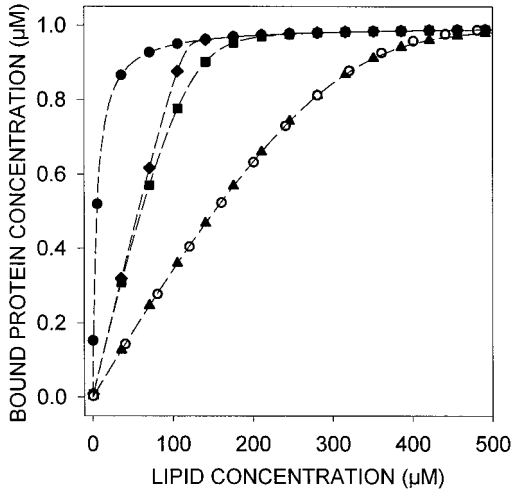

FIGURE 1 Simulations of the effect of protein self-association upon the binding of the protein to vesicles. The simulations were performed for an 8-mer with a K_X of 10^7 and a K_P of infinity (*closed circles*), 3.29×10^{-52} M^7 (*closed squares*), and 2.09×10^{-55} M^7 (*closed triangles*). In addition a simulation for a 30-mer and K_P of 2.94×10^{-212} M^{29} (*closed diamonds*) is shown. The open circles show the simulation defined by the closed triangles can also be fit by using a fixed value for K_X of 10^8 and a value of K_P, found by least squares, of 1.91×10^{-63} M^7.

Also shown in Fig. 1 is the simulated binding curve for a solute which aggregates to a 30-mer with a value of K_P set at 2.94×10^{-212} M^{29} (*closed diamonds*). This value of K_P was calculated from Eq. 2 assuming that the monomer concentration, in equilibrium with the 1 μM protein existing as a 30-mer, was also 0.05 μM. As can be seen, the binding curves for the 8-mer (*closed squares*) and 30-mer (*closed diamonds*) initially superimpose at low lipid concentrations, but the binding curve for the 30-mer maintains a longer linear portion until it sharply curves to join the curve obtained with the protein that does not aggregate (*closed circles*).

The reason for these "more linear" binding curves is seen in Fig. 2, where the corresponding concentrations calculated for protein in the unbound monomeric and aggregated forms in the same simulations are shown. The initial concentration of free monomeric protein (*dotted lines*) becomes progressively smaller as K_P is decreased, but then the concentration of monomeric protein declines only slowly from this initial value as the lipid concentration increases. In these latter situations it is the concentration of protein in the aggregated form (*solid lines*) which decreases as the lipid concentration increases. In essence, the aggregated protein is acting as a buffer for the monomeric protein and, as the latter remains relatively constant, the amount of bound protein is almost linearly proportional to the lipid concentration, while appreciable concentrations of aggregated protein exist. Once the concentration of aggregated protein approaches zero then a more hyperbolic binding curve is obtained. It should be noted that the very small size of K_P is a result of its 7th-power (or 29th-power for the 30-mer) dependence. As will be shown later, the free energy term associated with this equilibrium is for the incorporation of N molecules into the

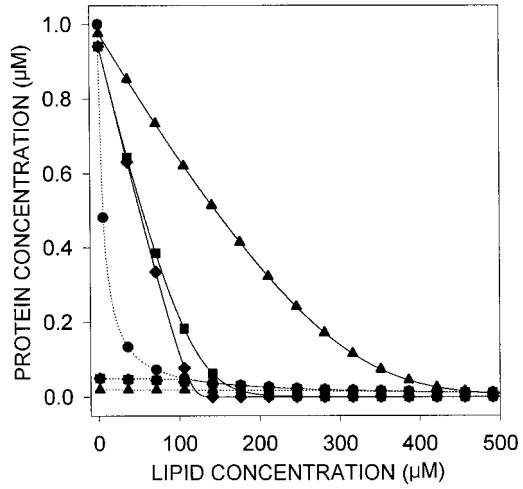

FIGURE 2 Simulations of the effect of protein self-association upon the concentration of unbound monomer and unbound aggregate in the presence of vesicles. The amounts of the unbound monomer (*dotted lines*) and unbound aggregate (*solid lines*) are from the same simulations shown in Fig. 1, and the same symbols are used.

N-mer. A more useful free energy value is for the incorporation of one molecule into the N-mer.

While these simulations generate binding curves that are obviously dependent upon all the input parameters, there is some interdependence between K_X and K_P which might, intuitively, prevent a unique fit to an experimental binding isotherm. In order to test this concern, the original BASIC program was modified so as to globally search for values of K_X and K_P which would best fit (by least squares) the simulated data in Fig. 1. These results are also shown in Fig. 1 where the data that resulted from the simulation with K_X and K_P of 10^7 and 2.09×10^{-55} M^7, respectively, were subjected to the fitting program with K_X fixed at 10^8 and N fixed at 8. The best fit was now obtained with $K_P = 1.91 \times 10^{-63}$ M^7 (*open circles*), which corresponds to an initial monomer concentration of 0.002 μM rather than the 0.02 μM used to generate the original simulated data. Given that these fits are to "ideal" data, it is apparent that a unique fit to "real" experimental data would be very difficult to obtain when both K_P and K_X are allowed to vary simultaneously. If a unique value for K_X is to be obtained, some independent estimation of K_P is needed (see below).

Analysis of experimental binding data obtained with cytochrome b_5

The simulated binding isotherms shown in Fig. 1 at non-infinite values of K_P are quite similar to published binding curves for b_5 (Leto and Holloway, 1979; Tretyachenko-Ladokhina et al., 1993; Taylor and Roseman, 1995). Such binding curves have been analyzed according to various models but, as far as we are aware, the self-association of the protein in the aqueous phase has not yet been incorporated into any of these models. As a test of the concerns expressed above, previously published fluorescence data for the binding of b_5 to lipid vesicles (Tretyachenko-Ladokhina et al., 1993) were subjected to the analysis described in the present report.

The original binding data for b_5 were obtained by fluorescence titration. This technique is often used to monitor the binding of solutes to membranes and depends upon the fluorescence enhancement that occurs as the solute moves from the aqueous phase into the hydrophobic environment of the bilayer. The overall fluorescence (F_T) of a sample containing monomer, aggregate, and vesicle-bound solute will be:

$$F_T = \frac{F_W \cdot [S]_W + F_N \cdot [S_N]_W + F_L \cdot [S]_L}{[S]_W + [S_N]_W + [S]_L} \qquad (3)$$

where F_W, F_N, and F_L are the fluorescence values of the solute when completely monomeric in aqueous solution, completely aggregated in aqueous solution, or completely bound to the vesicle, respectively. F_W and F_N can be determined once the aggregatory properties of the system, as described in Eq. 2, are known. F_L is difficult to obtain, unless binding is very tight, as light scattering will interfere with the fluorescence measurements at high lipid vesicle concentrations and full binding will of course only be approached asymptotically. Interestingly, the method of Eisinger and Flores (1985), which is very effective in compensating for light scattering effects, may not be valid here as this procedure uses a sample dilution technique to extrapolate the observed fluorescence to infinite dilution, and this dilution will perturb the various equilibria.

The situation with b_5 is actually less complicated than the general case described above. Previous studies have defined the fluorescence changes that occur upon aggregation and vesicle binding of b_5. The binding of b_5 to lipid vesicles is accompanied by a twofold increase in fluorescence quantum yield, whereas the conversion of octamer to monomer is accompanied by only a 3% increase in quantum yield, and only a 1.5 nm red-shift in the emission maximum (Leto and Holloway, 1979).

Equation 3, above, then simplifies to:

$$F_T = (F_W \cdot ([S]_T - [S]_L) + F_L \cdot [S]_L)/[S]_T \qquad (4)$$

where F_W and F_L are the fluorescence intensity values of b_5 in the absence of lipid and when fully bound to vesicles, respectively, and $[S]_T$ is the total amount of b_5. Although Eq. 4 is valid irrespective of the mechanism of binding of the b_5, we have shown previously that the species of b_5 which binds to lipid vesicles is the monomeric form (Krishnamachary et al., 1994; Leto and Holloway, 1979). In addition, fluorescence studies have shown that neither b_5 nor even its isolated membrane-binding domain self-associate in the membrane (Freire et al., 1983).

The difficulty in assigning a value for the fluorescence of the fully bound b_5 will markedly influence the data analysis. This can be appreciated from the simulation shown by the closed triangles in Fig. 1, where only 98% of the protein is

bound even at a lipid concentration of 500 μM. At this lipid concentration appreciable light scattering would be occurring and any attempt to increase the lipid concentration further would be counter-productive. If the fluorescence enhancement seen at 500 μM lipid were used to calculate $[S]_L$ at each lipid concentration, a large systematic error would be introduced. Instead of picking an almost arbitrary value for F_L and fitting the derived values of $[S]_L$, it was decided to fit the original experimental fluorescence data. In addition, in an effort to circumvent the difficulty of choosing an appropriate value for F_L, the computer program used for fitting the simulated binding isotherms, as described above, was modified to allow the maximum fluorescence (F_L) to also be varied as the experimental fluorescence data were fit. In this way it should be possible to extract values of K_X and K_P from any experimental fluorescence data without ever reaching full binding of the protein, with its associated light scattering problems. In practice, an initial value for F_L was chosen (1.75) which was well below the maximum fluorescence value experimentally observed (1.82) and this value was then incremented in a step-wise manner in order to determine the sensitivity of the fitting procedure to the value chosen for F_L. For each value of F_L, the values of K_P, K_X, and N were then varied to produce calculated values of F_T that gave the best fit, as judged by the "sum-of-the squares-of the-residuals," to the actual experimental fluorescence binding data. In this way no preconceived assumptions were made as to the values of N or F_L. The residuals of these fits are shown in Fig. 3 and indicate that the best overall fit is obtained with a 25-mer using a maximum fluorescence value of 1.86. From Fig. 3 it appears that better fits to the experimental data are obtained with values of N greater than 12 when the maximum fluorescence is set between 1.83 and 1.85. Fig. 4 shows the simulated binding curves for 8-, 12-, 18-, and 25-mers with the maximal fluorescence set to 1.83, 1.84, 1.85, and 1.86, respectively, together with the experimental data (*closed circles*). The values of K_X determined by these simulations ranged from 2.2×10^6 for a 25-mer to 4.0×10^6 for an 8-mer. If attempts are made to fit the data without considering aggregation, then the fit shown by the open circles was obtained, which corresponds to a K_X of 0.9×10^6. This latter fit had a sum-of-the-squares-of-the-residuals 10 times that of the other fits in the figure. The same fit as that indicated by the open circles was also obtained when the data were fit by a nonlinear least squares program (Johnson and Frasier, 1985).

Although these fits to the experimental data are very good, and are almost indistinguishable by eye, there was concern that the fitting routine selected both large aggregation numbers and high values for the maximum fluorescence, and thus the final value of K_P could be in error. Previous studies had indicated that the aggregation number for b_5 was 8; however, this value was obtained in a different buffer system and it was noted that the amount of monomeric b_5 decreased dramatically as the buffer concentration was increased slightly (Calabro et al., 1976). More recently we have found that the concentration of monomeric b_5 in equilibrium with aggregated protein was also influenced by the type of buffer and its pH. These latter observations could be rationalized by measurement of the conductivity of the buffers, whereupon it was found that a constant monomer concentration was seen at constant conductivity, irrespec-

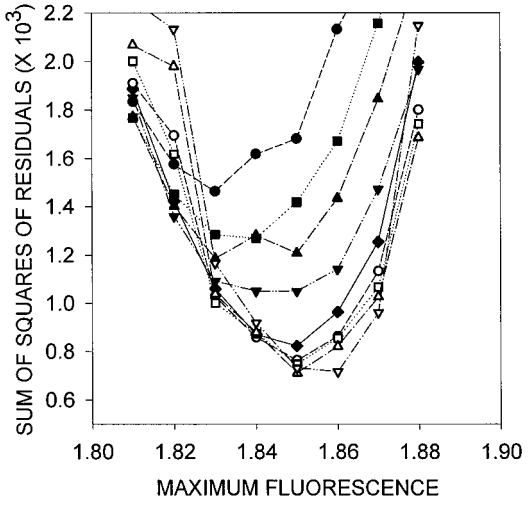

FIGURE 3 Fitting of experimental fluorescence data for b_5 binding to lipid vesicles. The iterative program was used to find the value of K_P which best fit the b_5 fluorescence binding isotherm as the value of K_L, the value of the maximum normalized experimental fluorescence enhancement upon binding, and the aggregation number (N) were varied. The data shown are for values for N of: 8, 9, 10, 12, 14 (*closed:* circle, square, triangle, inverted triangle, diamond), 16, 18, 20, and 25 (*open:* circle, square, triangle, inverted triangle).

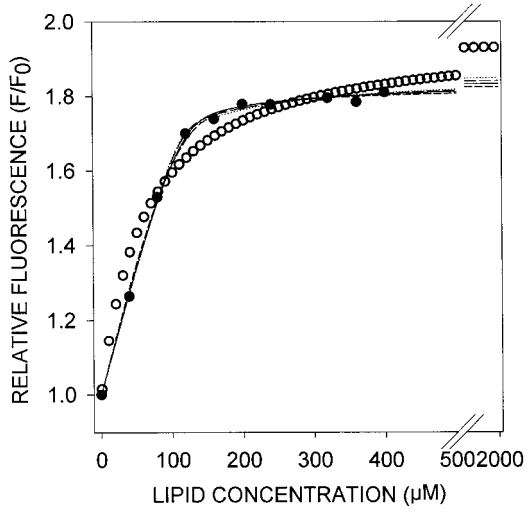

FIGURE 4 Global fitting of experimental fluorescence data for b_5 binding to lipid vesicles. The best fits to the experimental data (*closed circles*) for values of N of 8, 12, 18, and 25 were obtained from Fig. 3 and had values of K_X of 4.0, 3.2, 2.6, and 2.2×10^6. The binding curves simulated from these values are shown by the short dash, solid, medium dash, and dotted lines, respectively. The best fit to the experimental data assuming no aggregation was obtained with a K_X of 0.9×10^6 (*open circles*).

tive of buffer type and pH. This sensitivity to ionic strength is presumably due to the highly charged polar domain of the b_5 which contains 8 Asp, 10 Glu, 8 Lys, and 3 Arg residues. However, if slight increases in ionic strength of the buffer shielded these charges sufficiently to displace the equilibrium toward the aggregated form, then it could well allow more efficient packing and a larger aggregation number. In order to address these concerns about the previous, unrestricted, fitting procedure, where K_P, K_X, N, and F_L were all allowed to vary, it was decided to use gel filtration to obtain some estimation of K_P.

Analysis of cytochrome b_5 self-association by gel filtration

In an earlier publication the self-association of b_5 isolated from rabbit liver was examined by analytical ultracentrifugation and gel filtration (Calabro et al., 1976). In an aggregating system such as this, a complete thermodynamic analysis cannot be made by small-zone gel filtration (Ackers, 1970) but it can give some indication of the size of the aggregate and of the self-association equilibrium constant. If the self-association of a membrane protein is driven, as is its binding to membranes, by the hydrophobic effect, then it is likely that the aggregate contains many monomers and that the self-association is a cooperative process. There would be little shielding of the hydrophobic domain if lower-order aggregates were formed. This is what is seen with most detergents. Detergent micelles have quite defined aggregation numbers because of geometric considerations (Tanford, 1980) and gel filtration has been used to examine detergent micelle-monomer systems (Herries et al., 1964). When a solution containing detergent micelles is applied to an appropriate gel filtration column, the micelles elute near the void volume followed by a plateau corresponding to the concentration of monomer which at all times is in equilibrium with the micelles. If a self-aggregating solute is undergoing a cooperative aggregation, then depending upon the degree of aggregation, the monomer "plateau" seen upon gel filtration can be used to estimate K_P. With most detergents the aggregation number (N) is very large and the "plateau" concentration is the "critical micelle concentration" which is equal to $1/K_P$ (Tanford, 1980). However, even for much lower values of N, the monomer concentration is relatively independent of total solute concentration, as shown in Fig. 5. In Fig. 5 the dependencies of monomer concentration upon the total solute concentration were calculated from Eq. 2 with $[S]_W$ set at 0.05 μM, $[S]_T$ set at 1.7 μM, and values of N varied as 5, 6, 7, 8 (*dashed lines*), 9 (*dotted lines*), and 10, 12, 20, and 35 (*solid lines*). As can be seen, even with an 8-mer, by the time the concentration of solute reaches 2 to 4 μM the concentration of free monomer is relatively independent of further increases in total solute concentration. Hence, if an estimation of the concentration of monomer in a solution containing 2–4 μM of total protein can be obtained, then this can be used to calculate an approximate value of K_P by Eq. 2.

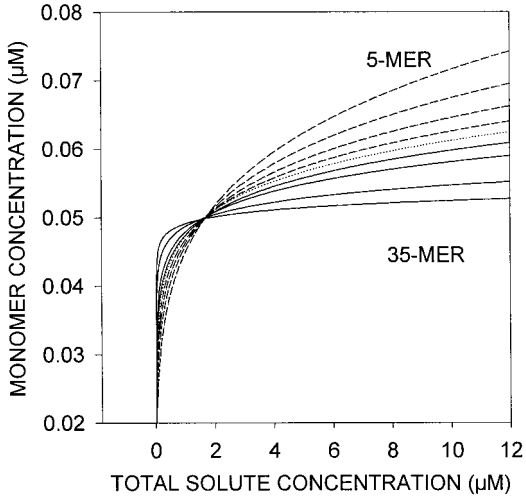

FIGURE 5 Relationship between monomer concentration and total solute concentration in a system with differing aggregation numbers. The equilibrium constant for the cooperative self-association of the different aggregates was calculated assuming the monomer and total solute concentrations were 0.05 and 1.7 μM, respectively. The curves, with decreasing monomer concentration, represent 5-, 6-, 7-, and 8-mer (*dashed lines*), 9-mer (*dotted line*), and 10-, 12-, 20-, and 35-mer (*solid lines*).

Gel filtration upon Sephadex G200 was performed with b_5 using the same buffer as used previously in the binding experiments. The results are shown in Fig. 6 and it can be seen in the inset of Fig. 6 that irrespective of the size of the sample applied and a 20-fold variation in concentration of the peak eluted in the void volume, the concentration of b_5 eluting between 1700 and 1900 s never exceeds ~0.05 μM and is relatively constant with time. This region corresponds

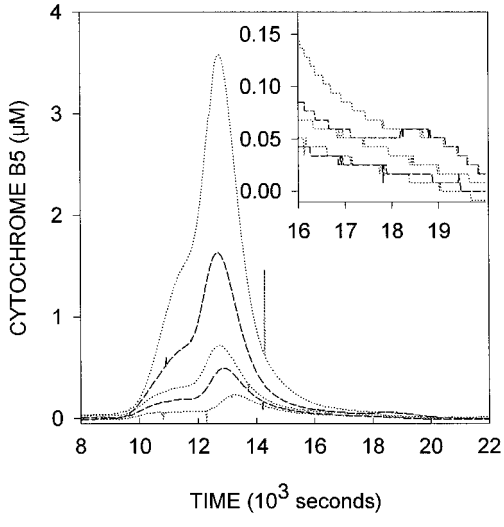

FIGURE 6 Gel filtration of cytochrome b_5. Samples of b_5 in 40 μl of HEPES buffer were mixed with 10 μl 5M KCl to ensure the protein was all aggregated and were then applied to the gel filtration column equilibrated with HEPES buffer at 25°C. The absorbance of the eluate was monitored at 412 nm. The samples applied contained 10, 5, 2.5, 1.9, or 1.25 nmol of b_5 from highest to lowest eluate profile. The line types in the inset correspond to the ones in the full figure.

to the "plateau" region seen in the gel filtration of detergents, but is less defined because of the relative Stokes' radii of the monomeric and aggregated protein species. The concentrations of the b_5 in the sample (0.05 ml) applied to the column ranged from 25 to 200 μM, whereas the peak concentration of b_5 in the aggregate eluting from the column ranged from 0.2 to 3.7 μM, ~100-fold lower because of dilution during the gel filtration. Most critically, despite a doubling of concentration of the applied sample from 100 to 200 μM, which resulted in a corresponding near-doubling of the concentration of the aggregate peak, from 1.6 to 3.6 μM, there was no increase in the concentration of the included peak, 0.05 μM. This is a characteristic of an aggregating system. In addition, the concentration of the aggregate peak in these two experiments, 1.6 and 3.6 μM, is very close to the total concentration, 1.7 μM, used in the previous experiments where the binding of b5 to vesicles was followed by fluorescence (Tretyachenko-Ladokhina et al., 1993). Hence, it will be assumed, based on these gel filtration data, that the concentration of monomeric protein in equilibrium with aggregated protein at the start of those previous fluorescence experiments was 0.05 μM. Given the lack of information on the precise value of the aggregation number, the possibility of nonspecific interactions between the protein and the gel-filtration matrix under the low ionic strength conditions (10 mM HEPES-0.1 mM EDTA (pH 7.3 at 25°C), conductivity 504 μS), and the precision with which these low absorbances (<0.01 OD) can be measured, it is felt that this value is a reasonable estimate of the maximum monomer concentration present in a solution of 1.7 μM b_5.

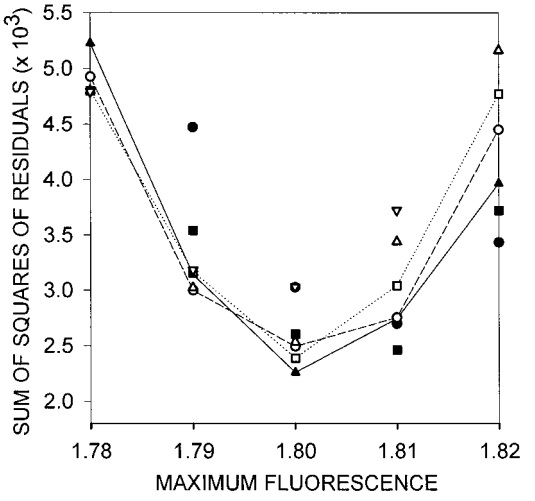

FIGURE 7 Fitting of experimental fluorescence data for b_5 binding to lipid vesicles. The value of K_P was fixed at the value calculated from the gel filtration data and the iterative program was used to find the value of K_X which best fit the b_5 fluorescence binding isotherm, as the value of the maximum normalized experimental fluorescence enhancement upon binding and the aggregation number were varied. The data shown are for values of N of 5, 6, 7 (*closed:* circle, square, triangle), 8, 9, 10, and 12 (*open:* circle, square, triangle, inverted triangle), and lines are drawn through the data for values of N of 7 (*solid line*), 8 (*dashed line*), and 9 (*dotted line*).

Analysis of the binding data for cytochrome b_5 using the K_P obtained by gel filtration

The original binding experiments with b_5 (Tretyachenko-Ladokhina et al., 1993) were performed at a protein concentration of 1.7 μM. Based upon the gel filtration data it is assumed that this solution should contain monomeric protein at 0.05 μM and K_P can be calculated from Eq. 2 [$(N \times 0.05^N)/1.65$]. The values of K_P for aggregation numbers of 5, 6, 7, 8, 9, 10, and 12 are: 9.5×10^{-31} M^4, 5.7×10^{-38} M^5, 3.3×10^{-45} M^6, 1.9×10^{-52} M^7, 1.1×10^{-59} M^8, 5.9×10^{-67} M^9, and 1.8×10^{-81} M^{11}, respectively. These values of K_P, and the respective N, were inserted as fixed values into the simulations, again allowing K_X and the maximum fluorescence to vary for each K_P and N pair. The residuals of the fits obtained are shown in Fig. 7 and indicate that the best fits are now obtained with N = 7–10 with the maximum fluorescence set at 1.80. Under these circumstances, the 12-mer or higher aggregates do not provide as good a fit. The binding curves for the simulations for all N are shown in Fig. 8 with the maximum fluorescence set to 1.80. In this plot, compared to Fig. 4 where the fluorescence value of 1.84 gave the best fits, the higher fluorescence values are not given as much weight which, given the risk of light-scattering contributions at higher lipid concentrations, is probably advisable. The values of K_X which produced these simulations are 19.5, 19, 18, 18, 17.2, 17, and 17×10^6 for the 5-, 6-, 7-, 8-, 9-, 10-, and 12-mer, respectively.

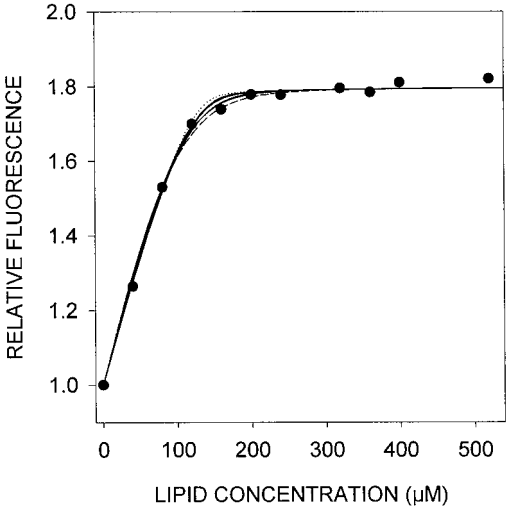

FIGURE 8 Restricted fitting of experimental fluorescence data for b_5 binding to lipid vesicles. The best fits to the experimental data (*closed circles*) obtained from Fig. 7, with K_P fixed at the value calculated from the monomer concentration of 0.05 μM obtained from gel filtration data, for 5-mer and $K_X = 19.5 \times 10^6$ (*dashed line*), 6-mer and $K_X = 19.0 \times 10^6$, 7-mer and $K_X = 18.0 \times 10^6$, 8-mer and $K_X = 18.0 \times 10^6$, 9-mer and $K_X = 17.2 \times 10^6$, 10-mer and $K_X = 17.0 \times 10^6$ (*solid lines*), and 12-mer and $K_X = 17.0 \times 10^6$ (*dotted line*) are shown.

The experimental binding data for native b_5 can thus be well fit with a molar partition coefficient, K_X, of 18×10^6 with the additional constraint that the monomer which binds to the vesicle is in equilibrium with an aggregated (N = 7–10) form of the protein with a dissociation constant, K_P, ranging from 3.3×10^{-45} M^6 for the 7-mer to 5.9×10^{-67} M^9 for the 10-mer. There is no necessity for inclusion of negative cooperativity or of a "binding-to-sites," although these may indeed be part of the mechanism of the binding of b_5 to vesicles.

When these binding data were analyzed previously (Tretyachenko-Ladokhina et al., 1993) the binding curves were analyzed by a method in which the normal binding process is modified to account for the "number of lipids in the binding site" (Hille et al., 1981), but aggregation of the protein in the aqueous phase was not taken into account. This model considers the interaction of the solute and vesicle as an equilibrium reaction where each solute binds to Z molecules of lipid, and renders those lipid molecules unavailable for binding further molecules of solute (Hille et al., 1981):

$$K_L = [S]_W \cdot ([L] - Z \cdot [S]_L)/[S]_L \quad (5)$$

where K_L is the dissociation constant of the reaction (Hille et al., 1981) and, as before, $[S]_L$ and $[S]_W$ are the molar concentrations of monomeric solute in the vesicle fraction and aqueous fraction (bound and free), and [L] and [W] are the molar concentrations of lipid and water, respectively. When the binding data were refitted with this model, also allowing the fluorescence maximum to vary in a stepwise manner, the results shown in Table 1 were obtained. Very good fits can be obtained with this model, actually better than with the partition model (see Fig. 7), and the best fit was obtained with a fluorescence maximum of 1.81, a K_L of 4.4 μM, and 65 lipids in the binding site. When the aggregation of the protein was also taken into consideration and the best fits to the experimental data were obtained with each K_P and N pair, as K_L, fluorescence maximum, and the number of lipids in the binding site were varied, it was found that there was little change in the value of K_L, the fit deteriorated, and became worse as more lipids were included in the binding site. This lack of improvement of fit with the inclusion of aggregation is to be expected as the amount of protein bound at low lipid levels is limited primarily by the number of "sites" and only secondarily by the available monomeric protein.

TABLE 1 Fitting of binding isotherm for cytochrome b_5 assuming no aggregation

Final fluorescence	K_L (μM)	Number of lipids	Sum of squares of residuals
1.79	3	65	2.75×10^{-3}
1.80	3.3	66	1.54×10^{-3}
1.81	4.4	65	1.12×10^{-3}
1.82	6.0	63	1.18×10^{-3}
1.83	7.2	63	1.55×10^{-3}
1.84	9.0	62	2.31×10^{-3}

Experimental data previously published for the binding of b_5 to POPC SUV (Tretyachenko-Ladokhina et al., 1993) were subjected to the least-squares fitting program, using the "binding-to-sites" model with the assumption that no aggregation was occurring. The values of K_L, which gave the lowest sum-of-the-squares of the residuals for all combinations of the value of the maximum normalized experimental fluorescence enhancement upon binding and the number of lipids in the binding site, were determined and the overall best fit for each fluorescence value is listed.

A very complete study of the binding of bovine b_5 to various lipids was performed by Taylor and Roseman (1995). They used a Scatchard analysis in studies of the binding of bovine b_5 to POPC SUV in 20 mM Tris-acetate, 100 mM NaCl, and obtained a value of 17 for the lipids in the binding site and an association constant (per lipid) of 2.2×10^4 M^{-1}, which corresponds to a dissociation constant of 45 μM. The aggregation properties of bovine b_5 have not been examined in detail and we have not generated binding isotherms with the rabbit b_5 in the presence of high salt, so a comparison between the two systems is not possible. However, from other measurements (manuscript in preparation), we find binding of rabbit b_5 to POPC SUV is 40% tighter in high salt.

If the self-association of hydrophobic ligands, such as membrane proteins, is a common phenomenon, then it is instructive to examine the advantage of obtaining an estimate, however approximate, of K_P. In the present system the estimated concentration of monomer from the gel filtration experiment was 0.05 μM. As a test of the sensitivity of the fitting procedure, values of monomer were also set at 0.033 μM and 0.075 μM. These were again used to calculate K_P for different aggregation numbers. The values of K_X and N which gave the best fits for the three chosen values of monomer, 0.033, 0.05, and 0.075 μM, were 28×10^6 with a 6-mer, 18×10^6 with a 7-mer, and 12×10^6 with a 7-mer. These values of K_X can be compared to the values derived by the global fitting where K_P was not restricted: $K_X = 2.2 \times 10^6$ with a 25-mer having a $K_P = 5.5 \times 10^{-155}$ M^{24}. The gel filtration procedure would clearly have excluded the latter fit as this K_P corresponds to a monomer concentration of 0.3 μM.

With many hydrophobic ligands the aggregate may not be of defined size or may be very large; in these cases also, an estimation of the concentration of monomer would still be of value. The simulations in Fig. 1 showed that identical initial slopes are seen for both an 8-mer and a 30-mer, each in equilibrium with identical monomer concentrations. The major difference between the two binding curves is the "sharpness" of the curvature at high levels of binding. These linear portions of the curves are occurring where the monomer concentration is constant, hence Eq. 1 reduces to: $K_X = ([S]_L/[L])/([\text{monomer}]/[W])$. If a binding experiment gives a nonhyperbolic binding curve with an appreciable linear portion, the slope of the linear portion multiplied by 55.5/[monomer] is K_X. If K_X is being determined by a single partition measurement between two phases or between vesicles and buffer, then, provided the true monomer concentration can be determined by gel filtration, by centrifuga-

tion, or by ultrafiltration, then K_X can be calculated. All of this assumes that only the monomer partitions into the hydrophobic phase, as did the detailed analysis above, and as has been shown specifically with the b_5 system.

These K_X and K_P values can be used to determine the ΔG of the two processes. Two groups have recently made determinations of the ΔG of membrane partitioning and both noted the difficulty of choosing the correct K values for the calculation (Wimley and White, 1993; Jones and Gierasch, 1994). We have, for simplicity, used the mole fraction partition coefficient (K_X), with $\Delta G = -RT \ln K_X$, which gives a value for ΔG of -9.9 kcal mol^{-1}. The values of N which produced the best fits (Fig. 6) were in the range of 7–10. If the self-association is assumed to be to an 8-mer, then by analogy with detergent micellization (Tanford, 1980), for each monomer binding to the 8-mer the average $\Delta G = (RT/8) \ln (1.9 \times 10^{-52}) = -8.8$ kcal mol^{-1}. Within the errors of the estimation of K_P by gel filtration and the fitting procedures, it appears that the ΔG for membrane association is slightly more negative than the ΔG for self-association, as is observed qualitatively in binding studies.

The techniques used here can be used in other systems even where the solute aggregate does not have a reasonably defined aggregation number; however, we think that this report should also serve as a cautionary tale. The very property, hydrophobicity, which causes peptide domains to insert into bilayers must also tend to cause the peptide domains to self-associate. That this self-association occurs so proficiently with b_5 may be due to the two-domain amphipathic nature of the protein, but it should be suspected with any hydrophobic peptide or solute. Although self-association in the membrane is much more relevant to biological processes such as pore formation (Schwarz et al., 1986; Arkin et al., 1995) or helix-helix interaction (Bormann and Engelman, 1992), and self-association in the aqueous phase would probably never occur under physiological concentrations, the latter phenomenon can severely complicate the interpretation of experimental binding data.

The authors thank Dr. Thomas E. Thompson for his critical reading of the manuscript.

This work was supported by the Thomas F. Jeffress and Kate Miller Jeffress Memorial Trust. Dr. N. Başaran was supported by NATO TUBITAK.

REFERENCES

Ackers, G. K. 1970. Analytical gel chromatography of proteins. *Adv. Protein Chem.* 24:343–446.

Arkin, I. T., M. Rothman, C. F. Ludlam, S. Aimoto, D. M. Engelman, K. J. Rothschild, and S. O. Smith. 1995. Structural model of the phospholamban ion channel complex in phospholipid membranes. *J. Mol. Biol.* 248:824–834.

Başaran, N., A. W. Steggles, and P. W. Holloway. 1996. The binding of native and mutant cytochrome b_5 to vesicles containing unsaturated lipids. *Biophys. J.* 70:264.

Bormann, B. J., and D. M. Engelman. 1992. Intramembrane helix-helix association in oligomerization and transmembrane signaling. *Annu. Rev. Biophys. Biomol. Struct.* 21:223–242.

Calabro, M. A., J. T. Katz, and P. W. Holloway. 1976. The self-association of cytochrome b_5 in aqueous solution: gel filtration and ultracentrifugational studies. *J. Biol. Chem.* 251:2113–2118.

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

Eisinger, J., and J. Flores. 1985. Fluorometry of turbid and absorbent samples and the membrane fluidity of intact erythrocytes. *Biophys. J.* 48:77–84.

Freire, E., T. C. Markello, C. Rigell, and P. W. Holloway. 1983. Calorimetric and fluorescence characterization of the interactions between cytochrome b_5 and phosphatidylcholine bilayers. *Biochemistry*. 22: 1675–1680.

Herries, D. G., W. Bishop, and F. M. Richards. 1964. The partitioning of solutes between micellar and aqueous phases: measurement by gel filtration and effect on the kinetics of some bimolecular reactions. *J. Phys. Chem.* 68:1842–1852.

Hille, J. D. R., G. M. Donne-Op den Kelder, G. H. de Haas, and M. R. Edmond. 1981. Physicochemical studies on the interaction of pancreatic phospholipase A_2 with a micellar substrate analogue. *Biochemistry*. 20:4068–4073.

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

Johnson, M. L., and S. G. Frasier. 1985. Nonlinear least-squares analysis. *Methods Enzymol.* 117:301–342.

Jones, J. D., and L. M. Gierasch. 1994. Effect of charged residue substitutions on the thermodynamics of signal peptide-lipid interactions for the *Escherichia coli* LamB signal sequence. *Biophys. J.* 67:1546–1561.

Krishnamachary, N., F. A. Stephenson, A. W. Steggles, and P. W. Holloway. 1994. Stopped-flow fluorescence studies of the interaction of a mutant form of cytochrome b_5 with lipid vesicles. *J. Fluorescence*. 4:227–233.

Ladokhin, A. S., L. Wang, A. W. Steggles, and P. W. Holloway. 1991. Fluorescence study of a mutant cytochrome b_5 with a single tryptophan in the membrane-binding domain. *Biochemistry*. 30:10200–10206.

Leto, T. L., and P. W. Holloway. 1979. Mechanism of cytochrome b_5 binding to phosphatidylcholine vesicles. *J. Biol. Chem.* 254:5015–5019.

Rizzo, V., S. Stankowski, and G. Schwarz. 1987. Alamethicin incorporation in lipid bilayers: a thermodynamic study. *Biochemistry*. 26: 2751–2759.

Schwarz, G., S. Stankowski, and V. Rizzo. 1986. Thermodynamic analysis of incorporation and aggregation in a membrane: application to the pore-forming peptide alamethicin. *Biochim. Biophys. Acta*. 861: 141–151.

Spatz, L., and P. Strittmatter. 1971. A form of cytochrome b_5 that contains an additional hydrophobic sequence of 40 amino acid residues. *Proc. Natl. Acad. Sci. USA*. 68:1042–1046.

Tanford, C. 1980. The hydrophobic effect. Wiley, New York.

Taylor, K. M. P., and M. A. Roseman. 1995. Effect of cholesterol, fatty acyl-chain composition, and bilayer curvature on the interaction of cytochrome b_5 with liposomes of phosphatidylcholines. *Biochemistry*. 34:3841–3850.

Tretyachenko-Ladokhina, V. G., A. S. Ladokhin, L. Wang, A. W. Steggles, and P. W. Holloway. 1993. Amino acid substitutions in the membrane-binding domain of cytochrome b_5 alter its membrane-binding properties. *Biochim. Biophys. Acta.* 1153:163–169.

Wimley, W. C., and S. H. White. 1992. Partitioning of tryptophan side-chain analogs between water and cyclohexane. *Biochemistry*. 31: 12813–12818.

Wimley, W. C., and S. H. White. 1993. Membrane partitioning: distinguishing bilayer effects from the hydrophobic effect. *Biochemistry*. 32:6307–6312.

Wimley, W. C., and S. H. White. 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Struct. Biol.* 3:842–848.