Eur Biophys J (1985) 12: 143–151

European Biophysics Journal

Protein lateral distribution in lipid bilayer membranes

Applications to ESR studies*

D. J. Laidlaw and D. A. Pink**

Theoretical Physics Institute, St. Francis Xavier University, Antigonish, Nova Scotia, Canada B2G 1CO

Received October 11, 1984/Accepted April 15, 1985

Abstract. We analyse recent ESR measurements on Ca^{2+} ATPase and Myelin proteolipid apoprotein reconstituted in phosphatidylcholine bilayer membranes. Our intention is to discover whether the measurements indicate significant protein-protein repulsive or attractive interactions. In order to do so we have studied a model of a lipid bilayer membrane containing transbilayer proteins. It represents the proteins by hexagons moving on a triangular lattice interacting via an energy E_0 which can be attractive, repulsive or zero. The last-named represents the "random" case studied earlier. We find that all of the Ca^{2+} ATPase data is best described either by the "random" model or, possibly, by one in which there is a small repulsive interaction, but not by the "annulus" model or one in which there is always at least one layer of lipid chains between every pair of proteins. We find that all of the Myelin PLA data is best described by a "random" distribution of hexamers and not by an "annulus" model of hexamers. We suggest measurements that can be done in order to unambiguously settle the question of whether these systems are best described by a "random"-type model or an "annulus"-type model.

Key words: Lipids, proteins, electron spin resonance, computer simulation

Introduction

There has been some debate about the interpretation of data on lipid bilayers containing integral proteins at $T > T_c$ obtained from steady state fluorescence polarization measurements using the probe DPH, and ESR measurements using nitroxide or proxyl spin labels. The early ESR measurements by Jost et al. (1973) were interpreted as showing that each integral protein was surrounded by an unbroken "annulus" of lipids. That is, there was a fixed stoichiometric ratio of "boundary lipids" to integral protein. Subsequent ESR data was interpreted in terms of this model (e.g. Jost and Griffith 1978; Knowles et al. 1979; Marsh and Watts 1982). This interpretation is based on the fitting of the protein concentration-dependence of the "immobilized component fraction", $f_i(c)$, to the function, $f_i^A(c) = Mc/(1-c)$. Here M is the maximum number of lipid molecules that can fit around an isolated transbilayer protein and c is the mole fraction of proteins in the bilayer.

In analyzing DPH data on cytochrome oxidase in DMPC bilayers, as well as the ESR data for the same system, Hoffmann et al. (1981) suggested that the best fit to both sets of data was a curve different from that predicted by the "annulus" model. They concluded that the proteins were distributed "randomly" in the plane of the bilayer so that protein-protein "contacts" could occur. This implies that there is not an unbroken lipid annulus around each protein and that there is no fixed stoichiometric ratio of "boundary lipids" to integral protein. A consequence of this was that the magnitude of M is larger than had been reported by earlier studies. Recently, Pink et al. (1984) applied this analysis to Ca^{2+} ATPase in bilayer membranes and reached similar conclusions, in contrast to the interpretation of Silvius et al. (1984) using the same data.

The physical interactions leading to these two models might be envisaged as follows: The "annulus" model represents the case in which protein hydrophobic sections are always at least two lipid chain layers apart, possibly due to repulsion on the

Abbreviations: PC, phosphatidylcholine; DMPC, dimyristoyl PC; DPPC, dipalmitoyl PC; EYPC, egg yolk PC; 14-PCSL, 1-acyl-2-[14-(4,4-dimethyloxazolidine-N-oxyl)steroyl]-sn-glycero-3-phosphocholine; DPH, 1,6-diphenyl-1,3,5-hexatriene; PLA, proteolipid apoprotein; ESR, electron spin resonance; T_c, temperature of main lipid phase transition.

* Work supported by NSERC of Canada

** To whom offprint requests should be addressed

part of overhanging or charged hydrophilic sections. The "random" model represents the case in which such repulsion is either absent or negligible. Two questions then arise: *(i)* Is it possible that the existing ESR data might be best fitted by a model in which some protein-protein repulsion exists but which is not so "hard" as to create a lipid "annulus" around each protein?[1] *(ii)* If some proteins might exhibit mutual repulsion, might others exhibit mutual attractive (albeit, possibly weak) interactions which could give rise to protein aggregation even for $T > T_c$?

In this paper we shall present model calculations and compare them with recent data on Ca^{2+} ATPase and Myelin proteolipid apoprotein. In the next section we shall discuss how to model such a system and conclude that computer simulations will most easily give the information required. We shall briefly describe the simulation method, and in the following section present and discuss the results. In the Appendix we present evidence that our simulation methods are correctly doing what we have described.

Theory

Models of protein lateral distribution in lipid bilayers

We want to answer the question: Given a lipid bilayer containing a concentration, c, of an integral protein, together with a very small concentration of spinlabelled lipids. What average fraction of labelled lipids are adjacent to the proteins? One way to calculate this is to perform a molecular dynamics simulation of at least half of the bilayer. The 3-dimensional structure of the lipid chains will be included together with that of the hydrophobic section of the proteins. By including the (known) interactions between the various hydrophobic constituents, one can calculate, not only the information sought, but also some details of molecular vibrations, relaxation times and static order parameters. There are at least two difficulties to be overcome: the structure of the proteins of interest are not known in sufficient detail, and simulations would have to be performed on very large systems for very long times in order to get reliable data. To represent the protein by a cylinder or other structureless object, without being able to assign surface structure to it, while retaining the space-filling details of lipid hydrocarbon chain structure, would seem inappropriate. It therefore becomes reasonable to discard the details of the lipid hydrocarbon chain structure and represent one half of the bilayer by a plane with shapes representing the average lipid chain and protein cross-sectional areas. These shapes can then move in the plane, thereby modelling the large-scale lateral motion of lipids and proteins, with their detailed conformational structure represented by one average shape.

The next question is: what shapes best represent average lipid and protein cross-sectional areas? One model is to represent lipid molecules by "hard" ellipses and proteins by "hard" discs or other shapes. Such "hard repulsive core" models have the property that two neighboring objects cannot interpenetrate each other and so can touch only at points. When such objects of two different sizes are mixed, one expects a number of "holes" (areas not filled by the objects) because of their packing properties.

A better model might be to represent the constituents of the bilayer by "soft" ellipses, discs or other shapes. Here the repulsive interaction between two neighbouring objects would not rise to infinity at their point of contact, but would rise to some finite value thus allowing some interpenetration. This model might better describe the packing of lipid chains by leaving fewer holes so as to better model a relatively impermeable bilayer at temperatures not close to the main lipid transition temperature.

A simple model which takes into account the properties of "soft core" lipid chains has been described (Pink et al. 1984). It considers the protein-lipid bilayer as a two-dimensional "gas" of proteins, moving in a "sea" of lipids, and subjected to random thermal forces. The model represents half of the bilayer by a triangular lattice. The cross-sectional area of proteins are represented by hexagons whose centers and vertices occupy lattice sites. Each hexagon occupies a number, n_H, of lattice sites and two hexagons may not occupy sites in common. Hexagons are thus adjacent when they are separated by one lattice constant. Sites not occupied by hexagons are assumed to be occupied by (labelled or unlabelled) lipid hydrocarbon chains. An isolated hexagon has M sites adjacent to it, and M represents the number of lipid molecules which can fit around an isolated trans-bilayer protein in both halves of the bilayer, with

$$n_H = 1 + M(M-6)/12 . \quad (1)$$

1 After this paper was submitted we learned (D. Chapman, private communication) that just such a modification of the "annulus" model has been proposed. It is suggested that some integral proteins possess either a complete unbroken annulus of lipids, i.e. the "annulus" model (so that the distance of closest approach of two proteins is two annulus' thickness), or can approach each other so as to share their annuli (so that their closest approach is one annulus thickness). The simulations performed here study the second case in the limit that $E_0 \to \infty$ (below)

The protein concentration, c, is defined to be

$$c = N_H/(N_H+N_c),\ N_c+n_H N_H = N_s\,, \tag{2}$$

where N_s is the total number of sites on the triangular lattice, N_H is the number of hexagons and N_c is the number of lipid chain sites.

In order to better understand what follows, let us briefly review the "annulus" and "random" models.

As far as the interpretation of ESR data is concerned, the "annulus" model assumes that each integral protein "immobilizes" M lipid molecules independently of the lipid to protein ration, L/P, or that a total of M lipid molecules can completely surround each protein at all protein concentrations, c. This "immobilized boundary layer" is the lipid "annulus". The fraction of the lipids so immobilized, $f_i^A(c)$, is thus proportional to the value of P/L so that, if labelled lipids are distributed like unlabelled lipids, then

$$f_i^A(c) = M(P/L) = Mc/(1-c)\,. \tag{3}$$

The "annulus" model thus assigns to pairs of proteins an infinitely repulsive second nearest neighbour interaction so that they can approach only as closely as third neighbours.

The "random" model assumes that, for some $T > T_c$, the proteins are "randomly" distributed in the plane of the bilayer. The "immobilized" ESR spectrum is assumed to arise from those (labelled) lipids which are adjacent to at least one integral protein (which does not necessarily imply that unlabelled lipids are immobilized). Labelled lipids are assumed to be distributed like unlabelled lipids, and the fraction of labelled lipids adjacent to at least one integral protein, thereby giving rise to the "immobilized" spectrum, is,

$$f_i(c) = 1-(1-c)^M + \Delta f_i(c)\,. \tag{4}$$

$\Delta f_i(c)$ is the fraction due to correlations not included in $(1-c)^M$ and it has been shown by computer simulations using hexagons on (finite-sized) triangular lattices, that $\Delta f_i(c)$ is less than $\sim 2.3\%$ of $1-(1-c)^M$ for a wide range of M and c (Pink et al. 1984). The "random" model has zero interaction between proteins so that they are free to approach as closely as nearest neighbours. Brotherus et al. (1981) considered the case wherein the labelled lipids are not distributed like the unlabelled lipids, and this can easily be included into modifications of (3) (Brotherus et al. 1981) or (4) (Pink, unpublished results).

Here we shall study cases different from these two extremes. We shall consider that two proteins can interact when they are nearest neighbours and that this interaction can either be repulsive (so enabling us to comment on the proposals in Footnote 1, p. 144) or attractive.

We model hexagon-hexagon (i.e., protein-protein) interactions as follows: It is clear that each lattice site is "connected" to each of its six nearest neighbour sites by a bond equal in length to the lattice constant. Two adjacent hexagons will be connected by a number, n_{bb}, of such bonds. We shall take the interaction between these hexagons to be Kn_{bb}, where each bond contributes K. The dimensionless constant K is the energy per bond, E_0, divided by k_BT, where k_B is Boltzmann's constant and T is the absolute temperature. If K, i.e., E_0, is positive, then the interaction is repulsive, while if it is negative, then the interaction is attractive. Figure 1 shows two adjacent hexagons with $n_{bb} = 5$. Here $n_H = 19$ so that $M = 18$.

In the case where $E_0 = 0$, the "random" case, we were able to write down an analytical expression for the fraction of sites adjacent to at least one hexagon (Hoffmann et al. 1981), and subsequently showed by computer simulation on finite-sized lattices (Pink et al. 1984) that the expression was an excellent approximation for all protein concentrations of interest. In order to calculate an analytical expression for $f_i(c)$ when $E_0 \neq 0$ we would need to consider multiple hexagon-hexagon correlations. In order to study $f_i(c)$ in these cases we shall use computer simulations because recent work (Gupta et al. 1984, and references therein) suggests that the superposition approximation used in calculating analytical expressions for 3-particle correlations in dense fluids may be unsatisfactory.

Simulation method

We have used the procedure described by Pink et al. (1984) modified to perform a Monte Carlo simula-

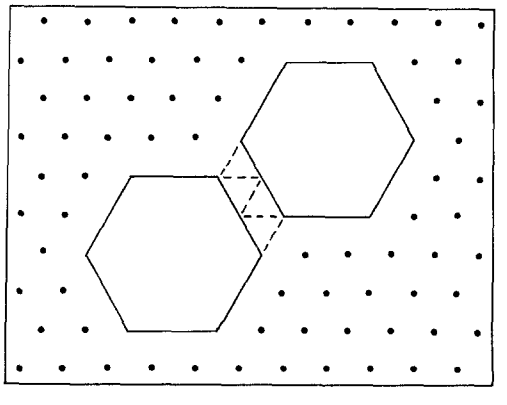

Fig. 1. The model of one half of a lipid bilayer. Sites represent lipid hydrocarbon chains and hexagons represent the cross-section of a transbilayer integral protein. Here $M = 18$, $n_H = 19$ and five hexagon-hexagon bonds, $n_{bb} = 5$, between two adjacent hexagons are shown as dashed lines

tion using Kawasaki dynamics (Kawasaki 1972; for a simple review and description of this procedure see Pink 1984) in order to calculate quantities characterizing equilibrium lateral distributions of hexagons on the lattice.

The simulation proceeds as follows: *(i)* Hexagons are distributed randomly (except possibly at very high concentrations) so that they do not overlap, on a triangular lattice of N_s sites possessing periodic boundary conditions. *(ii)* Each hexagon is visited once, and only once, in one "step". The order in which they are visited is determined by a random number generator and varies from step to step. When a hexagon is visited, one of the six directions in which it can be moved by one lattice constant on the lattice, is selected randomly. If the hexagon can be so moved without occupying sites already occupied by another hexagon, then the Kawasaki procedure is followed. If movement would involve occupying sites occupied by another hexagon, then the hexagon cannot move, the Kawasaki procedure is not followed, and another hexagon is visited. *(iii)* The Kawasaki procedure is as follows: Let the hexagon posses a total of n bonds with other adjacent hexagons before its move, and would possess n' bonds with adjacent hexagons if the move were carried out. The corresponding interactions are Kn and Kn'. Define

$$\Delta K = K(n' - n) . \tag{5}$$

If $\Delta K \leqq 0$ then the move is carried out. If $\Delta K > 0$, then select a random number, R, lying between 0 and 1. If $R \leqq \exp(-\Delta K)$ then the move is carried out. If $R > \exp(-\Delta K)$ then the move is not carried out. *(iv)* When all hexagons have been visited once, the step has been completed. *(v)* The procedure *(ii)* is carried out for a sufficiently large number of steps to "initialize" the system. After initialization, the procedure *(ii)* is carried out for a number of steps large enough to enable the calculation of averages.

At each step we calculated *(a)* the fraction of lipid chain sites, $f_l(M, c, K)$, adjacent to at least one hexagon. When $K = 0$ this quantity becomes equal to $f_l(c)$ of Eq. (4). *(b)* The fraction of such sites, $p_1(M, c, K)$, adjacent to only one hexagon. *(c)* The number of mutually adjacent hexagons, $N_{HH}(M, c, K)$, and the number of hexagon-hexagon bonds, $N_{bb}(M, c, K)$. We then averaged these quantities over the number of steps performed after initialization.

The quantity *(a)* should give the fraction of the ESR spectrum due to the "immobilized" component in the case of nitroxide or similar spin labels attached sufficiently near the terminal methyl group of saturated lipid hydrocarbon chains. The other three quantities describe the lateral distribution of hexagons on the lattice and, by implication, the lateral distribution of intrinsic proteins in the plane of the bilayer.

Results

In view of the current interest in Ca^{2+} ATPase (Silvius et al. 1984) and Myelin PLA (Brophy et al. 1984) we studied the case of $M = 24$. In the case of Ca^{2+} ATPase in delipidated sarcoplasmic reticulum membranes, as well as reconstituted in DPPC and in EYPC, studies have been carried out of steady state fluorescence polarization using DPH (Gomez-Fernandez et al. 1980) and ESR using nitroxide- or proxyl-labelled PC probes (Jost and Griffith 1978; Thomas et al. 1982; Silvius et al. 1984). The values of M reported range from 17 (Marsh and Watts 1982) or 22 (Silvius et al. 1984) per Ca^{2+} ATPase monomer for analyses using the "annulus" model, to 24, 25 and 28 per Ca^{2+} ATPase monomer for analyses using the "random" model (Pink et al. 1984). Myelin PLA has a molecular weight of between $\sim 25{,}000$ and $\sim 30{,}000$, and each isolated monomer should be able to accommodate a maximum of between ~ 24 to ~ 30 lipid molecules around it in both halves of the bilayer.

All results discussed in this section are the averages of two separate simulations, each averaged over 1,200 steps, except for the "random" case, $K = 0$, which involved a single simulation averaged over 1,200 steps. Simulations were carried out on triangular lattices of $N_s = 6{,}400$ sites. We consider that the results will not change substantially if larger lattices and larger numbers of steps are used.

Figure 2I shows $f_l(24, c, K)$ for $K = 0$, 0.5 and 50.0 (sets of points E, C and B). In addition, we have plotted $1 - (1 - c)^{24}$ as a solid line, D, ("random" model) and $24c/(1 - c)$ ("annulus" model) as a dashed line, A. Finite-lattice effects and errors due to the number of steps could yield error bars of the order of 2%, and possibly larger for concentrations greater than about 0.055 ($L/P \approx 17$). It can be seen that the simulations of $f_l(24, c, 0)$ are in good agreement with the curve $1 - (1 - c)^{24}$ unless $c \gtrsim 0.055$. A repulsive interaction of $K = 0.5$ per bond corresponds to a repulsive energy of $E_0 = 0.21 \times 10^{-13}$ ergs/bond at 30 °C. This should be compared to the energy of $\sim 0.45 \times 10^{-13}$ erg to excite one gauche conformer on an all-trans hydrocarbon chain (Flory 1969). A comparison of the data points C, with the solid line curve of $1 - (1 - c)^{24}$ or the data points D shows the effect of this repulsive energy. The effect is nearly at a maximum, however, because for $K = 50.0$ the data points B show little difference from those of C. In the case of B there will be, in general,

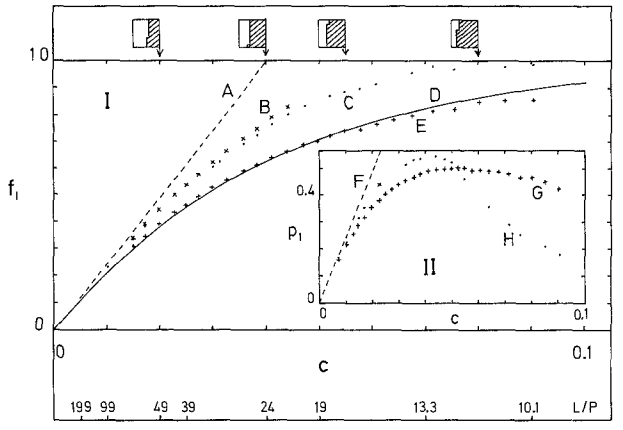

Fig. 2. Analytical (A, D, F) and computer simulation (B, C, E, G, H) results for **(I)** the fraction of lipid chain sites adjacent to at least one hexagon, $f_i(M,c,K)$, and **(II)** the fraction of chain sites adjacent to only one hexagon, $p_1(M,c,K)$. The case of repulsive or zero hexagon-hexagon interactions, K, with $M=24$. The hexagon concentration, c, is in mole fraction and L/P shows the corresponding lipid molecule to protein ratio. Simulations were performed on triangular lattices of 6,400 sites. B, C and H are results of two Monte Carlo runs each of 1,200 steps, while E and G are from one run of 1,200 steps. The errors for $c<0.055$ are less than $\sim 2\%$ while for $c \gtrsim 0.055$ they can be as large as $\sim 5\%$. **(I)** *A:* "Annulus" model $f_i^A = 24c/(1-c)$. *B:* $f_i(24,c,50.0)$. *C:* $f_i(24,c,0.5)$. *D:* $f_i(c) = 1-(1-c)^{24}$, "random" model. *E:* $f_i(24,c,0)$, "random" model. **(II)** *F:* "Annulus" model, $p_1^A = 24c/(1-c)$. *G:* $p_1(24,c,0)$, "random" model. *H:* $p_1(24,c,0.5)$

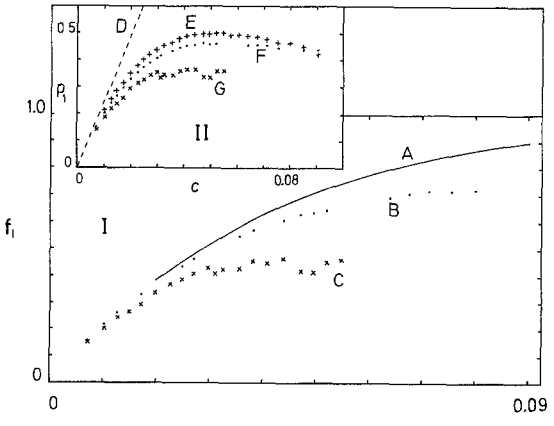

Fig. 3. Analytical (A, D) and computer simulation (B, C, E, F, G) results for **(I)** the fraction of lipid chain sites adjacent to at least one hexagon and **(II)** the fraction of chain sites adjacent to only one hexagon. The case of attractive or zero hexagon-hexagon interactions, K, with $M=24$. Simulations performed on triangular lattices of 6,400 sites. B, C, F, G are from two Monte Carlo runs of 1,200 steps each while E is from one such run. Hexagon concentration, c, is in mole fraction. Simulation errors are as in Fig. 2. **(I)** *A:* $f_i(c) = 1-(1-c)^{24}$, "random" model. *B:* $f_i(24,c,-0.1)$. *C:* $f_i(24,c,-0.19)$. **(II)** *D:* "Annulus" model, $p_1^A = 24c/1-c)$. *E:* $p_1(24,c,0)$ "random" model. *F:* $p_1(24,c,-0.1)$. *G:* $p_1(24,c,-0.19)$

at least one layer of lipid chain sites between every pair of hexagons for c as large as ~ 0.06. This can be seen in Fig. 2II which shows the probability that a chain site is adjacent to only one hexagon. The dashed curve F is for the "annulus" model – all chain sites adjacent to at least one hexagon are also adjacent to only one hexagon so that F and A are identical. Data points G are for the "random" model, $p_1(24,c,0)$, while points H are $p_1(24,c,0.5)$ for $K=0.5$. The fact that at higher concentrations very few chain sites are adjacent to only one hexagon shows that the hexagons are generally separated by at least one layer of chain sites unless the packing at high c forces hexagon-hexagon contact.

Figure 3I shows the corresponding results for attractive interactions. Curve A is $1-(1-c)^{24}$ while points B and C are $f_i(24,c,K)$ for $K=-0.1$ and $K=-0.19$ respectively. These values correspond to energies of $E_0 \approx -0.04 \times 10^{-13}$ erg and $E_0 \approx -0.08 \times 10^{-13}$ ergs respectively per hexagon-hexagon bond at 30 °C. The effect of the aggregation of the hexagons can be clearly seen in the very much smaller interface between the hexagon regions and the chain sites since it is the size of this interface which is measured by $f_i(M,c,K)$. Figure 3II shows the plots of $p_1(M,c,K)$ for the "annulus" model (D), the "random" model, $K=0$ (E), $K=-0.1$ (F) and $K=-0.19$ (G).

Figure 4I shows an analysis of recent ESR data, using a 14-proxyl spin label, obtained from EYPC bilayers reconstituted with Ca^{2+} ATPase at 25 °C (Silvius et al. 1984). The dashed curve, A, is a plot of $Mc/(1-c)$ for the "annulus" model with $M=22$ and represents the fit reported by Silvius et al. The solid line, C, is a plot of $1-(1-c)^M$, the result of the "random" model, with $M=28$. These points have been reported elsewhere (Pink et al. 1984). Our intention here is to see whether a better fit can be achieved by including a repulsive interaction between the hexagons. The data points X joined by the long-dashed curve B shows $f_i(24,c,0.5)$. For comparison $1-(1-c)^{24}$, dots D, is shown. Although for low concentrations $f_i(24,c,0.5)$ might fit the data marginally better than $1-(1-c)^{28}$, the latter appears to give a better fit at the two high-concentration points. It appears, then, that a "random" or, possibly, "near-random" model ($K<0.5$) fits all of the data better than does the "annulus" model.

One point is clear, however. In the region between $c \approx 0.0225$ and $c \approx 0.03$ further measurements such as those reported by Silvius et al. (1984) should be able to distinguish between the "random" model and the "annulus" model.

Figure 4II shows ESR results reported by Brophy et al. (1984) for Myelin PLA in DMPC at

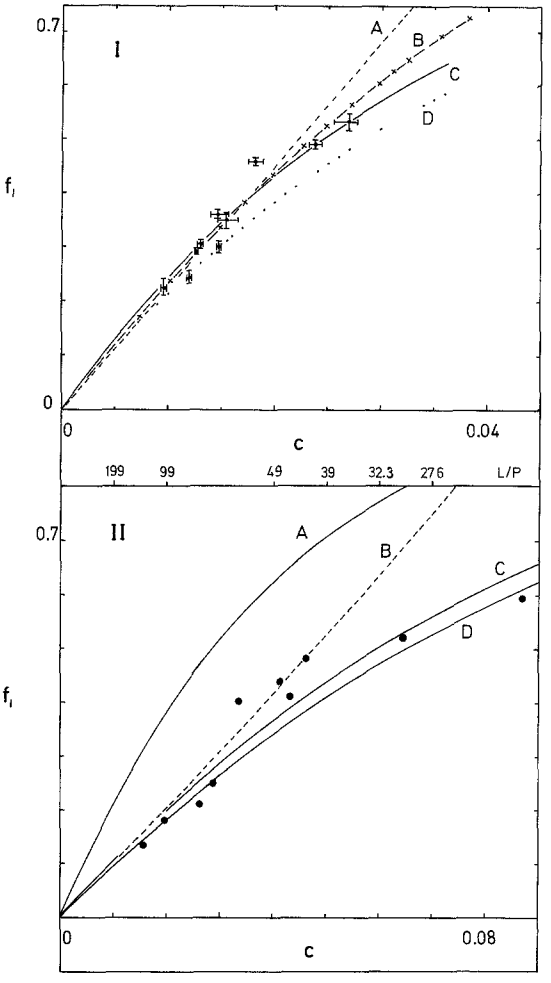

Fig. 4. The fraction of the ESR signal due to the "immobilized" component as a function of protein concentration, c (mole fraction), together with plots of the fraction of lipid chain sites adjacent to at least one hexagon. **(I)** Data points for Ca^{2+} ATPase in EYPC, ● (Silvius et al. 1984). *A:* "Annulus" model with $M = 22,22\,c/(1-c)$. *B:* $f_i(24, c, 0.5)$, simulations as in Fig. 2I with a repulsive interaction, $K = 0.5$ and $M = 24$. *C:* $f_i(c) = 1-(1-c)^{28}$, "random" model with $M = 28$ (Pink et al. 1984). *D:* $f_i(c) = 1-(1-c)^{24}$ "random" model with $M = 24$. **(II)** Data points for Myelin proteolipid apoprotein in DMPC at 30 °C, ● (Brophy et al. 1984). *A:* $f_i(c) = 1-(1-c)^{24}$ "random" model with $M = 24$. *B:* "Annulus" model for hexamers, $60\,c_6/(1-c_6)$. *C:* $1-(1-c_6)^{66}$, "random" model with $M = 66$ (11 lipids per monomer). *D:* $1-(1-c_6)^{60}$ "random" model with $M = 60$. $c_6 = c/(6-5\,c)$

30 °C using a 14-PCSL lipid as a probe. They have been replotted from their Fig. 3 and show the "immobilized" fraction as a function of protein concentration, c, in mole fraction. The slope at very low concentrations is ~ 8.7 which is clearly not in accord with a picture of protein monomers for which $M \approx 24$. It has been pointed out (e.g. Pink 1984) that if the monomers form long-lived oligomers made up of k monomers each, then the concentration, c, Eqs.

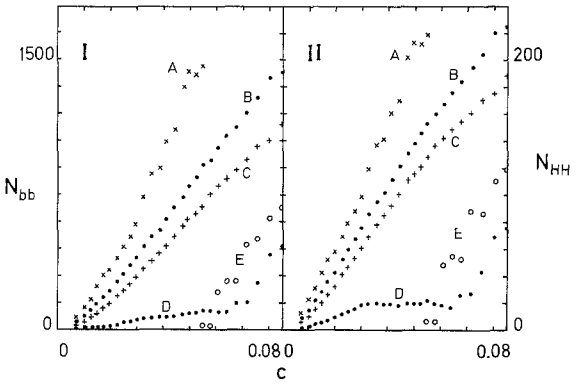

Fig. 5. Computer simulation results for the number of mutually adjacent hexagons, N_{HH} and hexagon-hexagon bonds, N_{bb}, (see fig. 1), on triangular lattices of 6,400 sites, with $M = 24$. **(I)** $N_{bb}(24, c, K)$. **(II)** $N_{HH}(24, c, K)$. *A:* $K = -0.19$. *B:* $K = -0.1$. *C:* $K = 0$, "random" model. *D:* $K = 0.5$. *E:* $K = 50.0$

(3) and (4) could be replaced by c_k where

$$c_k = c/(k-(k-1)\,c)\,. \tag{6}$$

Curve *A* shows $1-(1-c)^{24}$, the "random" model for $k = 1$ and $M = 24$, which clearly does not describe the data. The dashed curve *B* shows the prediction of the "annulus" model for $k = 6$ and $M = 60$ (10 lipid molecules per PLA monomer) as suggested by Brophy et al. (1984). The solid curves, *C* and *D*, show the predictions of the "random" model for $k = 6$, with $M = 66$ and $M = 60$ respectively. Thus, while we agree that the data is in accord with a picture of Myelin PLA forming hexamers in this system, we suggest that the data is better fitted by the "random" model for 10 or 11 lipid molecules per PLA monomer.

Figure 5 shows the number of hexagon-hexagon bonds, $N_{bb}(24, c, K)$, and the number of adjacent hexagons, $N_{HH}(24, c, K)$, averaged over 1,200 steps. It shows the cases of $K = -0.19$ (*A*), $K = -0.1$ (*B*), $K = 0$ ("random", *C*), $K = 0.5$ (*D*) and $K = 50.0$ (*E*). The corresponding numbers for the "annulus" model are zero up to $c = 0.04$, the maximum concentration which that model can achieve. Cases *A* and *B* show the effect of an attractive interaction which increases N_{bb} and N_{HH} above the "random" values. For $K = -0.19$ (*A*) a jog can be seen in both sets of points at $c \approx 0.015$ and a similar effect can be seen in points *C* of Fig. 3I at the same concentration. This is possibly due to changes involving aggregation of the hexagons. The points *D* and *E* for the cases of repulsive hexagon-hexagon interactions show that there are few hexagon-hexagon contacts in the case of $K = 0.5$, and that there are none ($N_{bb}(24, c, 50.0) \equiv 0$) in the case of $K = 50.0$, until the hexagon concentration exceeds $c \approx 0.055$. The

great variation in N_{HH} and N_{bb} as c increases above this concentration is probably due to the hexagon configurations traversing deep energy wells separated by high barriers.

Conclusions

We have studied a model of protein lateral distribution in a lipid bilayer at $T > T_c$ in order to calculate the protein concentration-dependence of the "immobilized" ESR signal due to lipid hydrocarbon chains labelled with nitroxide, or similar, spin labels sufficiently close to the terminal methyl group. We have been concerned with modelling cases where the proteins either repel or attract each other, previous papers (Hoffmann et al. 1981; Pink et al. 1984) having studied the case in which the proteins are "randomly" distributed.

We discussed ways of modelling this system and pointed out that a molecular dynamics simulation, in which the detailed structure of lipids and proteins is included, is the best model. In the absence of detailed information about protein structure in lipid bilayers, however, we have suggested that a model in which proteins and lipids are represented by "hard" shapes may not be a good one because of the existence of "holes" due to the packing properties of the components. A better model may be one in which proteins and lipids are represented by "soft" shapes so that fewer holes are formed. A similar result can be achieved by the model which we have used which represents the proteins by hexagons occupying the sites of a triangular lattice on which they can move. This model represents the protein-lipid bilayer as a 2-dimensional sheet in which the packing of proteins is important and the lipids (in this case displaying no specific binding to the proteins) simply fill the space between them.

We performed computer simulations on hexagons with $M = 24$ so that we could compare our results to those of ESR measurements involving Ca^{2+} ATPase and Myelin PLA. We showed how the ESR "immobilized" component signal fraction would depend upon protein concentration (in mole fraction) in cases where the proteins either attract or repel each other. Substantial differences with the "random" case were found for hexagon-hexagon bond energies of $\sim 0.21 \times 10^{-13}$ erg (repulsive) or ~ -0.04 and $\sim -0.08 \times 10^{-13}$ erg (attractive) at 30 °C. These energies are significantly smaller than the energy of $\sim 0.45 \times 10^{-13}$ erg required to excite a gauche conformer in a free all-trans hydrocarbon chain (Flory 1969).

We investigated the possibility that recent ESR data on Ca^{2+} ATPase reconstituted in EYPC bilayers (Silvius et al. 1984) shows evidence of direct protein-protein interactions. We conclude that either the "random" model or one in which there is a weak protein-protein repulsion ($K < 0.5$) might fit the published data equally well. Our results continue to suggest, however, that neither the "annulus" model, nor a model in which every pair of proteins has at least one lipid chain layer between them give as good a fit to the published data as either the "random" model or, possibly, one with a weak protein-protein repulsion.

We propose, however, that this question can be settled by a sufficient number of measurements with small "error bars" in the protein concentration range of ~ 0.0225 to ~ 0.03 where the lipid/protein monomer ratio ranges from ~ 45 to ~ 32.

We analysed recent ESR data on Myelin PLA in DMPC (Brophy et al. 1984) and suggest that while it is in accord with a model in which the proteins form long-lived (i.e. on a time-scale long compared to one in which the oligomers sample many equilibrium distributions in the plane of the bilayer) hexamers, the "random" model, with ~ 11 lipid molecules per monomer seems to give a better fit to all the data than does the "annulus" model. Again, more data, with small error bars, in the concentration range of $c \approx 0.03$ to $c \approx 0.06$ should be sufficient to distinguish between the two models.

Finally, in an Appendix we have shown that results for $f_i(c)$ in the case of the "random" model where no overlap of hexagons is allowed are quite different from the case in which "random" unrestricted overlapping of hexagons is permitted.[2] Computer simulations in both cases support the analytical expressions which we have derived, and the agreement suggests that we have made no subtle errors in the simulations.

Appendix

It is instructive to compare the results predicted by three ways of distributing hexagons on a triangular lattice in order to see how the number, f_i, the fraction of (non-hexagon) sites which are adjacent to at least one hexagon, depends upon the fractional area, f_A, of the lattice covered by hexagons.

2 It has been suggested (anonymous, private communication) that our computer simulations (Pink et al. 1984) contain gross errors so as to (erroneously) allow overlapping of hexagons. The purpose of the Appendix is to show that not only can we calculate exactly an expression for $f_i(c)$ in the case that overlapping is allowed, but also that our simulations of that case are in complete agreement with those equations, and that neither are remotely similar to the "random" case, Eq. (4), where no overlapping is allowed

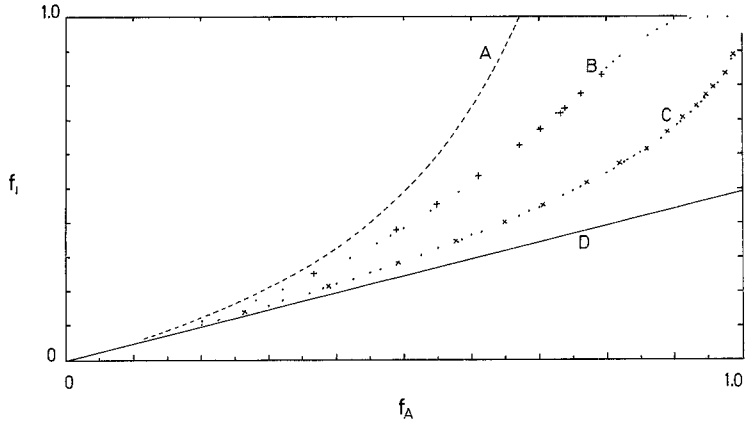

Fig. 6. The relation between the fraction of lipid chain sites adjacent to at least one hexagon, f_I, and the fraction of the total area covered by the hexagons, f_A, for $M=30$ ($n_H=61$). Simulations performed on triangular lattices of 10^4 sites. *A:* "Annulus" model (*dashes*). *B*: "Random" model. $f_I(c) = 1-(1-c)^M$ (*dots*) and simulations (+). *C:* "Random unrestricted overlap" model. $f_I^{RUO} \sim 1-(1-f_A^{RUO})^{M/n_H}$ (*dots*) and simulations (×)

In cases where hexagons are not allowed to overlap, i.e. cases where two hexagons may not occupy sites in common, the fraction of the lattice covered is

$$f_A = n_H N_H / N_s\,, \tag{A-1}$$

where N_H is the number of hexagons, n_H is the number of sites occupied by each hexagon and N_S is the total number of lattice sites. The hexagon concentration, $c = N_H/(N_H+(N_s-n_H N_H))$, and f_A are related by

$$c = f_A/(n_H-(n_H-1)f_A)\,. \tag{A-2}$$

Here we shall compare the predictions of the "annulus" and "random" models with that of a model in which "random" unrestricted hexagon overlaps (RUO) are allowed, i.e. two hexagons may occupy up to n_H sites in common.

(i) "Annulus" model. f_I^A is given by

$$f_I^A = \frac{Mc}{1-c} = \frac{Mf_A}{n_H(1-f_A)}\,. \tag{A-3}$$

(ii) "Random" model. At least to a very good approximation (Pink et al. 1984)

$$f_I = 1-(1-c)^M = 1-\left[\frac{n_H(1-f_A)}{n_H-(n_H-1)f_A}\right]^M. \tag{A-4}$$

(iii) "Random unrestricted overlap" model. For a lattice of N_s sites onto which N_H hexagons have been "randomly" distributed, allowing unrestricted overlapping to take place, we obtain

$$f_I^{RUO} = 1-\left(1-\frac{M}{N_s}\right)^{N_H}, \qquad f_A^{RUO} = 1-\left[1-\frac{1+M(M-6)/12}{N_s}\right]^{N_H}. \tag{A-5}$$

In the limit as N_s gets indefinitely large we obtain

$$f_I^{RUO} \sim 1-\exp(-MN_H/N_s)\,, \qquad f_A^{RUO} \sim 1-\exp(-n_H N_H/N_s) \tag{A-6}$$

so that

$$f_I^{RUO} \sim 1-(1-f_A^{RUO})^{M/n_H}. \tag{A-7}$$

Figure 6 shows the results of simulations to check (A-4) and (A-7) for the case of $M=30$ ($n_H=61$). The dashed curve, *A*, shows (A-3) with f_I^A reaching its maximum at $f_A \approx 0.67$ ($c \approx 0.032$). Points *B* are values of $1-(1-c)^M$ (dots) together with the results of computer simulation (+) on triangular lattices with $N_s = 10^4$ as described by Pink et al. (1984). Some new data is shown here. Points *C* show values of (A-7) (dots) together with computer simulation results (x) on lattices with $N_s = 10^4$. The straight line, *D*, shows the limiting slope of all these curves as $f_A \to 0$. The statistics for the simulation of the "random unrestricted overlap" model, *C*, are better than those of the "random" model, *B*, since about twice as many hexagons are needed in the former model to cover the same average area as covered by the non-overlapping hexagons in the latter model. In both cases the results of the simulations are in very good agreement with (A-4) and (A-7).

References

Brophy PH, Horvath LI, Marsh D (1984) Stoichiometry and specificity of lipid-protein interactions with myelin proteolipid protein studied by spin-label electron spin resonance. Biochemistry 23:860–865

Brotherus JR, Griffith OH, Brotherus MO, Jost PC, Silvius JR, Hokin LE (1981) Lipid-protein multiple binding equilibriums in membranes. Biochemistry 20:5261–5267

Flory PJ (1969) The statistical mechanics of chain molecules. John Wiley, New York

Gomez-Fernandez JC, Goni FM, Bach D, Restall CJ, Chapman D (1980) Protein-lipid interaction. Biophysical studies of (calcium-magnesium ion)-dependent-ATPase reconstituted systems. Biochim Biophys Acta 598:502–516

Gupta S, Haile JM, Steele WA (1984) Representation of static three-body correlations in dense fluids. Mol Phys 51: 675–696

Hoffmann W, Pink DA, Restall C, Chapman D (1981) Intrinsic molecules in fluid phospholipid bilayers. Fluorescence probe studies. Eur J Biochem 114:585–589

Jost PC, Griffith OH (1978) Lipid-protein interactions: Influence of integral membrane proteins on bilayer lipids. In: Agris PF (ed) Biomolecular structure and function. Academic Press, New York, pp 25–54

Jost PC, Griffith OH, Capaldi RA, and Vanderkooj G (1973) Evidence for boundary lipid in membranes. Proc Natl Acad Sci USA 70:480–484

Kawasaki K (1972) Kinetics of Ising models. In: Domb C, Green MS (eds) Phase transitions and critical phenomena. Academic Press, London, pp 443–501

Knowles PF, Watts A, Marsh D (1979) Spin-label studies of lipid immobilization in dimyristoylphosphatidyl-choline-substituted cytochrome oxidase. Biochemistry 18: 4480–4487

Marsh D, Watts A (1982) Spin labelling and lipid-protein interactions in membranes. In: Jost PC, Griffith OH (eds) Lipid-protein interactions, vol 2. John Wiley, New York, pp 53–126

Pink DA (1984) Theoretical studies of phospholipid bilayers and monolayers. Perturbing probes, monolayer phase transitions and computer simulations of lipid-protein bilayers. Can J Biochem 62:760–777

Pink DA, Chapman D, Laidlaw DJ, Wiedmer T (1984) Electron spin resonance and steady state fluorescence polarization studies of lipid bilayers containing integral proteins. Biochemistry 23:4051–4058

Silvius JR, McMillen DA, Saley ND, Jost PC, Griffith OH (1984) Competition between cholesterol and phosphatidylcholine for the hydrophobic surface of sarcoplasmic reticulum Ca^{2+}-ATPase. Biochemistry 23:538–547

Thomas DD, Bigelow DJ, Squier TS, Hidalgo C (1982) Rotational dynamics of protein and boundary lipid in sarcoplasmic reticulum membrane. Biophys J 37:217–225