A Monte Carlo Study of the Self-Assembly of Bacteriorhodopsin

Kamakshi Jagannathan, Rakwoo Chang, and Arun Yethiraj
Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706 USA

ABSTRACT Bacteriorhodopsin (BR) and specific lipid molecules self-assemble into a quasi two-dimensional lattice structure known as the purple membrane (PM). In the PM, BR molecules exist in a trimeric form with lipid molecules present in the space enclosed by each trimeric unit and in the inter-trimer space. These trimeric units, which have a roughly circular cross-section, are arranged in hexagonal patterns with long-ranged crystalline order. In this work, we investigate the self-assembly of BR in the PM via Monte Carlo simulations of a two-dimensional model of the membrane and proteins. The protein molecules are modeled as 120° sectors of a circle and the lipid molecules enter into the model through effective protein-protein interactions. The sectors cannot overlap with each other, and in addition to this excluded volume interaction there are site-site attractive interactions between specific points of the proteins to mimic interactions between helices on the proteins and lipid-induced interactions. At low values of the attractive well depth, the proteins are found in the monomeric form at all concentrations. At moderate and high values of the attractive well depth, trimers are formed as the concentration increases, and with a further increase in concentration the trimers organize into a hexagonal lattice. The interactions between the proteins and those induced by the intra-trimer lipids play an equally important role in the formation of trimers and the lattice. The lipids in the inter-trimer space cause the trimers to orient in a specific direction in the hexagonal crystal lattice.

INTRODUCTION

Bacteriorhodopsin (BR) is a transmembrane protein that uses light energy absorbed by its chromophore retinal to pump protons from the cytoplasm of archaea such as *Halobacterium salinarum* into the extracellular space (Oesterhelt and Stoeckenius, 1973). Research on BR has many facets, including interesting and important questions concerning the photochemistry of the retinal, protein folding, assembly into a naturally occurring hexagonal crystalline array, and conformational fluctuations (Lanyi, 2000; Booth, 2000). BR is exceptional in terms of its supramolecular organization and stability. BR is the main constituent of a two-dimensional (2D) crystalline lattice, the purple membrane (PM), which maintains structural and functional integrity under a wide range of pH, temperature, humidity and chemical environment (Oesterhelt et al., 1991). The PM is one of the best models for elucidating the interactions that are responsible for the assembly and stability of integral membrane protein complexes. In this article we present a simple model for the interactions and mechanism for the self-assembly of BR in the PM.

The structure of BR has been determined by high-resolution electron crystallography of isolated PM patches (Henderson and Unwin, 1975; Henderson et al., 1990; Grogorieff et al., 1996; Kimura et al., 1997; Essen et al., 1998). BR in three-dimensional (3D) crystals grown from lipidic cubic phase (Pebay-Peyroula and Rosenbusch, 2001; Lanyi and Luecke, 2001; Luecke et al., 1999; Belrhali et al., 1999) is arranged in hexagonally packed trimers similar to the native PM, indicating that these crystals are an accurate model of the PM. Each protein molecule consists of seven transmembrane α-helices surrounding the retinal chromophore and is arranged in 44 $\AA$ diameter trimeric units that pack in a hexagonal lattice. The protein trimers constitute ~75% of the PM by weight with the rest composed of lipids with fluid hydrocarbon chains. Lipid molecules are present between the trimers and in the space enclosed by each trimer. These detailed descriptions of the structure of BR and the associated lipids serve as the starting point for identifying the structural features of BR that determine its assembly in the PM.

The characteristic property of membrane proteins is their intimate contact with lipids. In the membrane, lipids and proteins form a tight seal that is prerequisite for the generation, maintenance, and controlled release of electrochemical potentials across a barrier, and for information transfer between the compartments. In the case of BR in the PM, the functional aspects, if any, are speculative; e.g., a steeper electric field gradient may be produced across the membrane than in the bulk lipid phase that might affect the proton pathway in BR (Essen et al., 1998). There are ~10 lipid molecules per molecule of BR in the PM (Grogorieff et al., 1996; Krebs and Isenbarger, 2000), comprising six or seven phospholipids, two or three sulfoglycolipids, and one squalene molecule. The lipid composition of the PM differs from that of the surrounding cytoplasmic membrane, and specific lipids are essential for the two-dimensional array formation of BR (Watts, 1995; Sternberg et al., 1993; Pebay-Peyroula and Rosenbusch, 2001; Weik et al., 1998; Fyfe et al., 2001). These lipids bind tightly to BR, stabilized in part by van der Waals packing of the hydrocarbon chains against the apolar surface of the protein. This high-affinity binding of the lipids to BR may be critical for lattice assembly and stability. Because specific lipids are present

Submitted March 11, 2002, and accepted for publication June 11, 2002

Address reprint requests to Arun Yethiraj, 1101 University Ave., Madison, WI 53706. Tel.: 608-262-0258; Fax: 608-262-9918; E-mail: yethiraj@chem.wisc.edu.

0006-3495/02/10/1902/15 $2.00

both in the inter-trimer space and in the space enclosed by each trimer, their role in the formation of trimers from BR monomers and in the association of the trimers in the 2D hexagonal lattice is of potential importance. There have been many mutational efforts to identify the structural features of BR that are critical for its assembly in the PM. The mutational studies concentrating on the protein-protein (BR-BR) interactions (Krebs et al., 1997; Isenbarger and Krebs, 1999) suggest that these interactions are important for the formation of the trimeric unit and for lattice stability. The presence of unique lipids in the lattice suggests, however, that BR-lipid interactions may also play a role in lattice stability (Krebs and Isenbarger, 2000). The thermal stability of the PM has been investigated by x-ray diffraction studies (Koltover et al., 1999) of hydrated PM films, which indicate a reversible melting of the lattice at ~75°C. A low degree of order is retained above this temperature, perhaps corresponding to the BR trimeric units. A denaturational transition at ~100°C occurs, consistent with the idea that the protein denatures as a trimer. These experiments can be complemented by theoretical and computational studies, where the relative importance of the protein-protein interactions and the lipid-induced interactions in the formation of BR trimers and the hexagonal lattice can be systematically and independently altered, thus providing unique insight into the mechanism of PM assembly and stability and also providing motivation for further experimental studies.

Molecular simulations can be performed on protein models ranging in complexity from detailed atomistic models to simple tractable coarse-grained models. There have been some molecular dynamics simulations of atomistic models of BR in the PM (Edholm et al., 1995; Baudry et al., 2001). These simulations analyze the structure and fluctuations of the protein and the lipids in the PM and the effect of water movement during BR's photocycle. Simulations of this nature cannot address the formation of the trimeric units of BR and the hexagonal lattice assembly because these processes occur on time scales several orders of magnitude larger than those accessible to molecular dynamics simulations. Stochastic Monte Carlo simulations (Sabra et al., 1998) have also been performed on a simple model consisting of two different lipid species, "annular" lipids and "neutral" lipids, and one protein species. The annular lipids are assumed to interact more strongly with the protein than the neutral lipids. Both the protein and the lipid species are modeled as disks with no internal structure. The simulations show that the formation of 2D arrays is facilitated by the presence of annular lipids. These earlier simulations (Sabra et al., 1998) do not address the self-assembly of the BR monomers to form trimers, nor do they elucidate the monomer-monomer interactions necessary for this assembly.

In this work we study the self-assembly of the BR trimers and the crystallization of the protein in the PM. We model the protein molecules as 120° sectors of a circle (Fig. 1). The apex of the sectors represents the lipid molecules present in the space enclosed by each trimer, and the midpoints of the two sides of the sectors represent the two helices of the protein that are expected to play a role in trimer formation. The lipids present in the interior of each trimer are referred to as "internal" lipids in this article. The lipids present in the inter-trimer space are represented by sites on the periphery of the sectors. These lipids are referred to as "external" lipids in this article. This model allows us to study the association of the BR monomers to form trimeric units in addition to the assembly of the trimers in hexagonal arrays. This was not possible in the disk model of proteins used by others (Sabra et al., 1998).

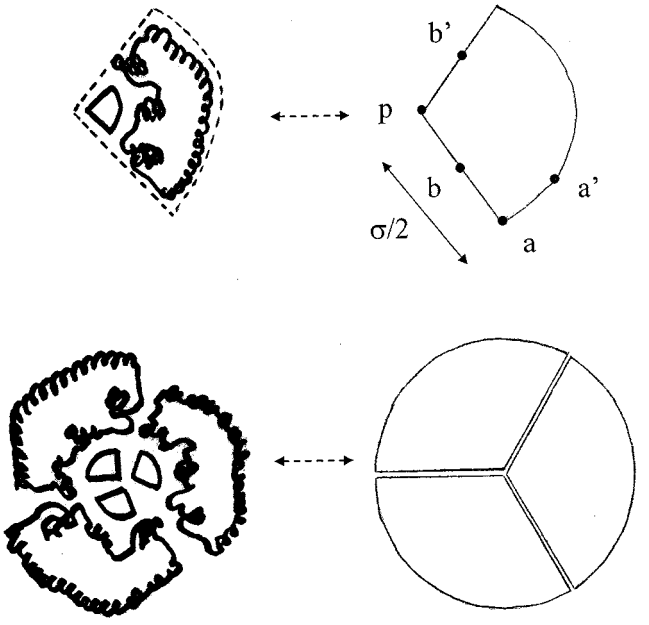

FIGURE 1 The 120° sector model used for the protein. The points b and b' represent the helices on the protein, p represents the internal lipids, and a and a' represent the external lipids. The trimeric unit formed by three monomers is represented by a circle formed by three sectors in this model.

We present grand canonical Monte Carlo simulations of BR using the 2D model. We are interested in studying the thermodynamics and possible phase transitions in trimer formation and lattice assembly, and in estimating the importance of the interactions between the transmembrane helices of adjacent protein molecules and the role of the lipid molecules in the formation of the trimeric units and the lattice. We consider protein-protein and lipid-lipid interactions (for both the internal lipids and the external lipids) and their effect on the phase behavior. We find that without any attractive interactions, i.e., only excluded volume interactions, neither trimers nor a lattice are formed at any concentration. Interactions between the internal lipids present in the space enclosed by each trimer or protein-protein interactions are essential for the formation of trimers and the hexagonal lattice, and these interactions have a similar effect on the phase behavior. At low concentrations the

protein is always in a monomeric form. For strong enough attractive interactions, trimers are formed as the concentration is increased, and when the concentration is increased further these trimers organize into a hexagonal lattice. Without interactions between sites outside the trimer, the hexagonal lattice thus formed lacks the specific orientational order seen in the PM. Incorporating interactions between the external lipids in combination with the interactions between the internal lipids gives a specific orientational order to the hexagonal lattice. We therefore conclude that the role of the protein-protein interactions and the lipid-lipid interactions of the internal lipids is to aid the formation of trimers and the hexagonal lattice, while the role of the external lipids is to provide a specific orientational order to the lattice.

The rest of the manuscript is organized as follows. In the next section we describe the model and the details of the interactions used. In subsequent sections we outline the simulation technique and present and discuss the results, summary, and conclusions.

MOLECULAR MODEL

We use a 2D model for the protein and membrane. The protein molecules are modeled as 120° sectors of a circle as shown in Fig. 1 and the membrane is treated as a continuum. The midpoints of the two sides of the sectors are labeled as b and b', and these represent the helices B and D in the protein and are believed to contribute important interactions in the self-assembly of BR (Krebs and Isenbarger, 2000). Although the precise location of phospholipids in the PM has not been determined, the location of sulfoglycolipid molecules have been determined in several studies. In vivo labeling and neutron diffraction studies of the PM have demonstrated sulfoglycolipid density at two locations per BR monomer: one in the space enclosed by the BR trimer and the other in the inter-trimer space (Weik et al., 1998). In our molecular model the apex of the sector is labeled as p, and this represents the lipids located in the interior of the BR trimer (internal lipids). The lipids present in the inter-trimer space (external lipids) are represented by two sites on the periphery of the sectors, labeled as a and a'. Interactions between the p sites and between the a and a' sites should be interpreted as lipid mediated protein-protein interactions. The length of each side of the sector is $\sigma/2$. In this work we define our length scale by setting $\sigma = 1$.

We incorporate excluded volume interactions between any two sectors, i.e., no part of any two sectors can overlap, and specific attractive interactions between the other sites. In different cases, we include attractions between b and b' sites, between p sites, and between a and a' sites. Square-well potentials are used for these site-site attractions, $V_{ij}(r)$, i.e.,

$$V_{ij}(r) = -\epsilon_{ij} \quad r \le \lambda$$
$$= 0 \qquad r > \lambda, \tag{1}$$

where ij can be pp, bb', or aa', ϵ_{ij} is the attractive well depth, λ is the range of the attractive potential, and r is the distance between the sites i and j on two different sectors. Most of the simulations are performed for $\lambda = 0.2$, to mimic a short-ranged attraction. The choice of λ, which defines the range of the attractive interaction, is important, and is discussed in connection with the results. We define a dimensionless parameter ϵ^*_{ij} as

$$\epsilon^*_{ij} = \frac{\epsilon_{ij}}{k_B T}, \tag{2}$$

where k_B is Boltzmann's constant and T is the absolute temperature. Changing ϵ^*_{ij} is equivalent to changing the attractive well depth ϵ_{ij}, at a fixed value of temperature. In this work we choose an energy scale by setting $k_B T = 1$.

SIMULATION METHOD AND PROPERTIES CALCULATED

The Monte Carlo simulations are a standard application of the Metropolis algorithm in the grand canonical ensemble (Allen and Tildesley, 1987; Frenkel and Smit, 1996), where the chemical potential, μ, area of the system, A, and temperature, T, are fixed. The system is open with respect to exchange of particles and energy. Successive configurations in the grand canonical simulations are generated by the following four moves: insertion, deletion, translation, and rotation of the individual sectors. Each new configuration generated by one of these moves is accepted with a probability chosen to ensure that the ensemble is sampled correctly (Allen and Tildesley, 1987; Frenkel and Smit, 1996). Simulations are performed for different values of ϵ^*_{ii} ranging from 1 to 8. The basis for choosing this range for ϵ^*_{ii} is as follows: $\epsilon^*_{ii} = 0$ corresponds to system consisting of hard sectors (with purely repulsive interactions in the form of excluded volume effects and no attractive interactions); $\epsilon^*_{ii} = 8$ corresponds to a system where the BR monomers exist as trimeric units even at the lowest value of the area fraction considered. The range $0 < \epsilon^*_{ii} < 8$ therefore spans the entire temperature range of interest. For each value of ϵ^*_{ii}, we consider different loadings of the system with the area fractions, ϕ, ranging from 0.1 to 0.68 over this range. In the PM, the BR trimers are densely packed in a hexagonal lattice. We consider the entire range of ϕ from a very dilute system to a fully loaded system in our simulations.

The simulations are performed in a square box of length 8σ in both the x and y directions. Periodic boundary conditions are imposed in both directions. The number of sectors, N, ranges from 30 for area fraction of 0.1 to 168 for area fraction of 0.68. Simulations are also performed for larger system sizes of 12σ and 16σ in both the x and y directions for a few select cases of ϵ^*_{ii} and area fraction ϕ. No significant finite size effects are observed.

The simulation proceeds in three stages: initial configuration generation, equilibration, and averaging. For the lowest concentrations, an empty box is used as an initial configuration. For simulations at a new state point, a final configuration from a previous simulation at a nearby state point is used as an initial configuration. For high area fractions, two different initial configurations, one disordered and one ordered, are used to ensure that the results do not depend on the initial configuration.

Starting with an initial configuration, successive configurations are generated by either attempting to move a sector or attempting an insertion or deletion of a sector. One of the four (translation, rotation, insertion, deletion) moves is chosen at random. In a translation move, the chosen sector is displaced by a random amount with a maximum displacement of 0.35σ. The percentage of accepted translation moves is ~40% for moderate values of ϵ^*_{ii} and ϕ and decreases as ϵ^*_{ii} and ϕ increase. For a rotation move, the chosen sector is rotated through a random angle with a maximum angular displacement of 20° (0.35 radians). Between 40 and 60% of attempted rotation moves are accepted. In the insertion move, a sector is inserted with the apex at a randomly chosen position and oriented in a randomly chosen angle. In a deletion move, a randomly selected sector is removed from the system. The percentage of the accepted insertion and deletion moves varies with the values of ϵ^*_{ii} and ϕ, and is ~30% for moderate values of ϵ^*_{ii} and ϕ. The percentage of accepted insertion and deletion moves decreases drastically at high values of ϵ^*_{ii} and ϕ. Approximately 2×10^8 attempted moves are required for equilibration of systems at low concentrations, with roughly twice as many attempted moves required for dense systems. Properties are calculated every 10^4 attempted moves after equilibration. The ensemble averaged properties are recorded every 5×10^7 attempted moves, for a total of 2.5×10^8 attempted moves. These five sets of ensemble averaged properties are used to obtain the final averaged properties and the statistical uncertainties.

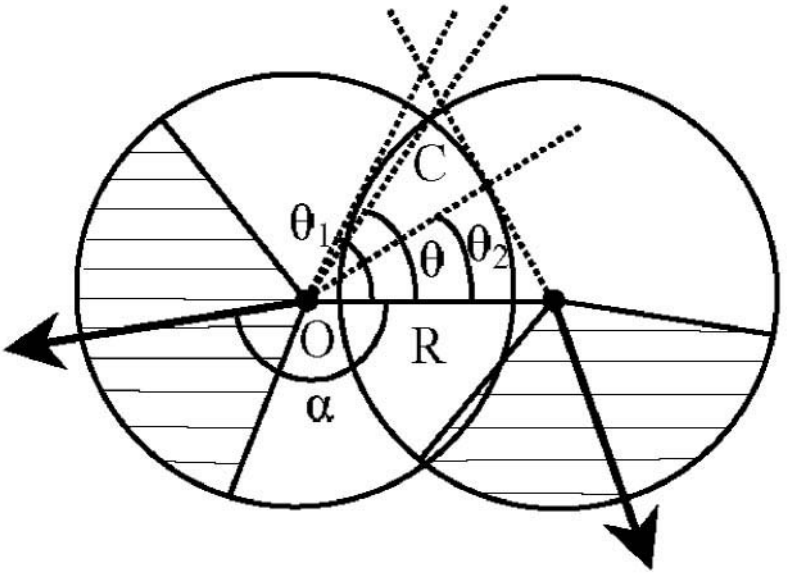

FIGURE 2 Conditions for overlap between two sectors. When the apexes are at distances less than σ, the sectors can overlap for certain values of the relative orientation.

The moves used in the simulations satisfy detailed balance, and this guarantees that the system samples state space in an unbiased fashion. The choice of maximum displacement (for rotation or translation) in a Monte Carlo scheme is arbitrary as long as the attempted displacement is chosen randomly. In this work, we choose the maximum displacement parameters with the aim of sampling configurational space efficiently. It is important to note that our choice of moves or maximum displacements will not affect the results of the simulations.

The most interesting part of the simulation is the checking of overlap between the sectors. In the case of hard spheres or hard disks, the check for overlap is very straightforward. If the distance between the centers of two spheres (disks) is greater than the hard sphere (disk) diameter, there is no overlap between them and there is an overlap of the two spheres (disks) otherwise. But in the case of the sectors, the check for overlap is complicated because the overlap of sectors depends on both the distance and the relative orientation (Fig. 2). The various cases depending upon the orientation of the two sectors have to be explicitly considered and checked for overlap. The algorithm for checking for overlap is presented in the Appendix.

The main focus of this study is on the structural and thermodynamic properties of the system. The radial distribution function, g_{pp} (r), is the probability, relative to an ideal gas, of finding a sector with its apex p at a distance r, given that there is another sector with its apex p at the origin. $g_{pp}(r)$ is defined by

$$g_{pp}(r) = \frac{A}{N^2}\left\langle \sum_i \sum_{j\neq i} \delta(r - r_{p_i p_j})\right\rangle, \qquad (3)$$

where A is the area of the system, N is the total number of sectors, and $r_{p_i p_j}$ is the distance between the apexes p of sectors i and j. The summations go over the total number of sectors, and the angle denote ensemble average. The standard histogram method is used to monitor $g_{pp}(r)$. To characterize the formation of BR trimers from the monomers, we calculate the number of sectors, n_b, whose apex point p is within a distance of 0.2σ from p of another sector. When a BR trimer is formed, each sector has two other sectors with the point p within a distance of 0.2σ and $n_b = 2$.

The formation of the hexagonal lattice is illustrated by a structure factor $S_{pp}(\mathbf{k})$, which corresponds to a diffraction pattern in an electron diffraction experiment. It is defined with respect to the apex point p as

$$S_{pp}(\mathbf{k}) = S_{pp}(k_x, k_y) = \left\langle \left| \sum_j e^{i\mathbf{r}p_j\cdot\mathbf{k}} \right|^2 \right\rangle, \qquad (4)$$

where $\mathbf{k}$ is the vector momentum transfer variable, $\mathbf{r}p_j$ is the 2D vector giving the position of the apex point p of a sector j, and the summation goes over the total number of sectors. If we consider a contour plot of the structure factor, plotted against k_x and k_y, there is a single peak at $k_x = k_y = 0$ for a random arrangement of the sectors. The intensity of this peak is equal to $\langle N^2\rangle$, as is evident from Eq. 4 for the structure factor. For a hexagonal assembly of the sectors, the contour plot has peaks at hexagonal positions with respect to $k_x = k_y = 0$. Because the trimers have a diameter of $\sim\sigma$, the first set of peaks at hexagonal positions can be expected at a distance of around $2\pi/\sigma$ for a hexagonal crystal (note that $\sigma = 1$ in this work).

In the grand canonical ensemble there are fluctuations in the total energy E and number of particles N. Thermodynamic properties can be calculated from these fluctuations. One such property is the specific heat C_v. In this ensemble, C_v is given as

$$C_v = \frac{3}{2}\langle N\rangle k_B + \frac{1}{k_B T^2}\left(\langle \delta U^2\rangle - \frac{\langle \delta U \delta N\rangle^2}{\langle \delta N^2\rangle}\right), \qquad (5)$$

where the first term is the ideal part of C_v, U is the potential energy contribution to the total energy, δU is the fluctuation in U, δU^2 is the fluctuation in U^2, and δN and δN^2 are fluctuations in N and N^2, respectively.

RESULTS AND DISCUSSION

Before we consider the results of our simulations, we briefly discuss the crystallization of hard spheres (McQuarrie, 1976) and hard disks. Hard spheres interact via purely repulsive forces, i.e., the only interaction is that two spheres

cannot overlap. Hard spheres undergo a first-order transition from a fluid to a solid phase at a reduced density, ρd^3, of 0.96 where ρ is the number of spheres per unit volume and d is the hard sphere diameter. This crystallization occurs solely because of entropic effects, i.e., for $\rho d^3 > 0.96$, the free energy of the crystal is lower than the free energy of the liquid.

The freezing of a system of hard disks is quite complex and is not completely understood. There is no conventional long-ranged order in these systems in the sense that the mean-square displacement of a disk from its ideal lattice position diverges for infinite systems, although the system does have long-ranged bond orientational order (Mermin and Wagner, 1966; Kosterlitz and Thouless, 1973; Halperin and Nelson, 1978; Young, 1979). For finite systems hard disks show a first-order phase transition from a liquid to a hexagonal crystal (Alder and Wainwright, 1962). A recent study suggests a weakly first-order transition at an area fraction of $\phi \approx 0.7$ (Jaster, 1999) in the thermodynamic limit. We are not concerned with the physics of strictly 2D systems which, although interesting, is not applicable to membranes, which are, after all, 3D objects. Although we do not observe any finite size effects in the range of box lengths we have investigated, our simulations are not expected to describe strictly 2D systems in the thermodynamic limit.

We have performed simulations of hard disks and hard sectors to investigate the liquid-solid transition in these systems. For hard disks we find a first-order transition from a liquid to a hexagonal crystal at an area fraction of $\phi = 0.56$ for a system with a fixed area of $64\sigma^2$. This transition from a liquid to a crystal occurs due to entropic effects: the disks in a crystal have greater translational entropy than in a liquid at the same density because they can sample space in a more efficient fashion. The system of hard sectors, however, has no phase transition from a liquid to a crystal. We have established this by starting with initial configurations in a crystalline state at very high area fractions; these configurations always melt into a disordered state. The absence of a crystal in systems of hard sectors is not obvious, although it is generally difficult to crystallize fluids that are composed of asymmetric or odd-shaped molecules. For example, atactic polymers are very difficult to crystallize. In our system, it is possible that the entropy loss due to confinement of the sectors upon trimerization is not compensated by the increase in entropy of trimers upon crystallization. We therefore conclude that the crystal formation of BR in the PM must be energetically favored, and investigate the effect of attractive interactions on the assembly.

We first consider sectors with p-p interactions, which represent the interactions between the internal lipids present in the space enclosed by each trimeric unit. As will be discussed shortly, we observe a number of phases as ϕ or ϵ^*_{pp} is changed. For clarity, we first summarize these phases and then describe their properties in greater detail. For low values of ϵ^*_{pp}, the sectors are disordered over the entire range of area fractions considered. At sufficiently high values of ϵ^*_{pp}, we see a sequence of structures as the area fraction is increased. At low area fractions, the sectors exist as a disordered fluid of monomers. As the area fraction is increased, some of the sectors assemble to form trimers. This is a smooth transition and the percentage of monomers that are in trimeric form increases smoothly as the area fraction is increased. At a certain value of the area fraction this liquid of trimers crystallizes into a hexagonal lattice of trimers.

The nature of these different "phases," i.e., disordered monomers, disordered trimers, and crystal, are ascertained by studying the pair correlation function, $g_{pp}(r)$, fraction of sectors, N_f, which are part of clusters of sizes $n_b = 0$, 1, and 2, and the structure factor, $S_{pp}(\mathbf{k})$. Fig. 3, *A–D* depicts structural properties of the system for four different cases. For an isotropic fluid of monomers, which occurs for $\epsilon^*_{pp} = 1$ and $\phi = 0.63$ (Fig. 3 *A*), the radial distribution function, $g_{pp}(r)$, has a peak at $r/\sigma = 0$ because the apex points p of two sectors can be very close to each other. In addition, there are peaks at integral multiples of 0.5σ, as the side of each sector measures 0.5σ. For large r, $g_{pp}(r) \rightarrow 1$, indicating that there is no long-ranged structure in the fluid. One of the inset plots shows the fraction of the total number of sectors, N_f, as a function of n_b. There are very few trimers in the system, i.e., the fraction of sectors with $n_b = 2$ is close to zero. The inset plot is a contour plot of the structure factor, $S_{pp}(\mathbf{k})$, plotted against k_x and k_y. The structure factor has a single peak at $k_x = k_y = 0$, indicating that there is no hexagonal assembly of the sectors. For this case, the fluid is clearly isotropic and composed primarily of monomeric sectors.

Under some conditions trimers exist, but the fluid is still a liquid. For $\epsilon^*_{pp} = 2.5$ and $\phi = 0.66$ (Fig. 3 *B*), $g_{pp}(r) \rightarrow 1$ for large r, indicating that there is no long-ranged order. This is also evident from the single peak at $k_x = k_y = 0$ in the structure factor, but the fraction of sectors plotted against n_b shows a significant number of sectors with $n_b = 2$, indicating a significant number of BR timers. This is also reflected in the $g_{pp}(r)$, where the peak at $r = 0.5\sigma$ has a much lower value than the peak at $r = \sigma$, because once trimers are formed the probability of finding the point p of a sector at a distance of 0.5σ from the point p of another sector decreases.

Once the molecules are in predominantly the trimeric state, hexagonal crystals of these trimeric units appear. For $\epsilon^*_{pp} = 4$ and $\phi = 0.64$ (Fig. 3 *C*), some long-ranged order is seen in $g_{pp}(r)$ for large r, and is reflected in faint peaks in the structure factor at hexagonal positions around $k_x = k_y = 0$. The fraction of sectors with $n_b = 2$ is also high, indicating the predominance of trimers. The peak at $r = 0.5\sigma$ in $g_{pp}(r)$ has almost vanished, consistent with the fact that all the sectors are part of trimers. For a more dense system with

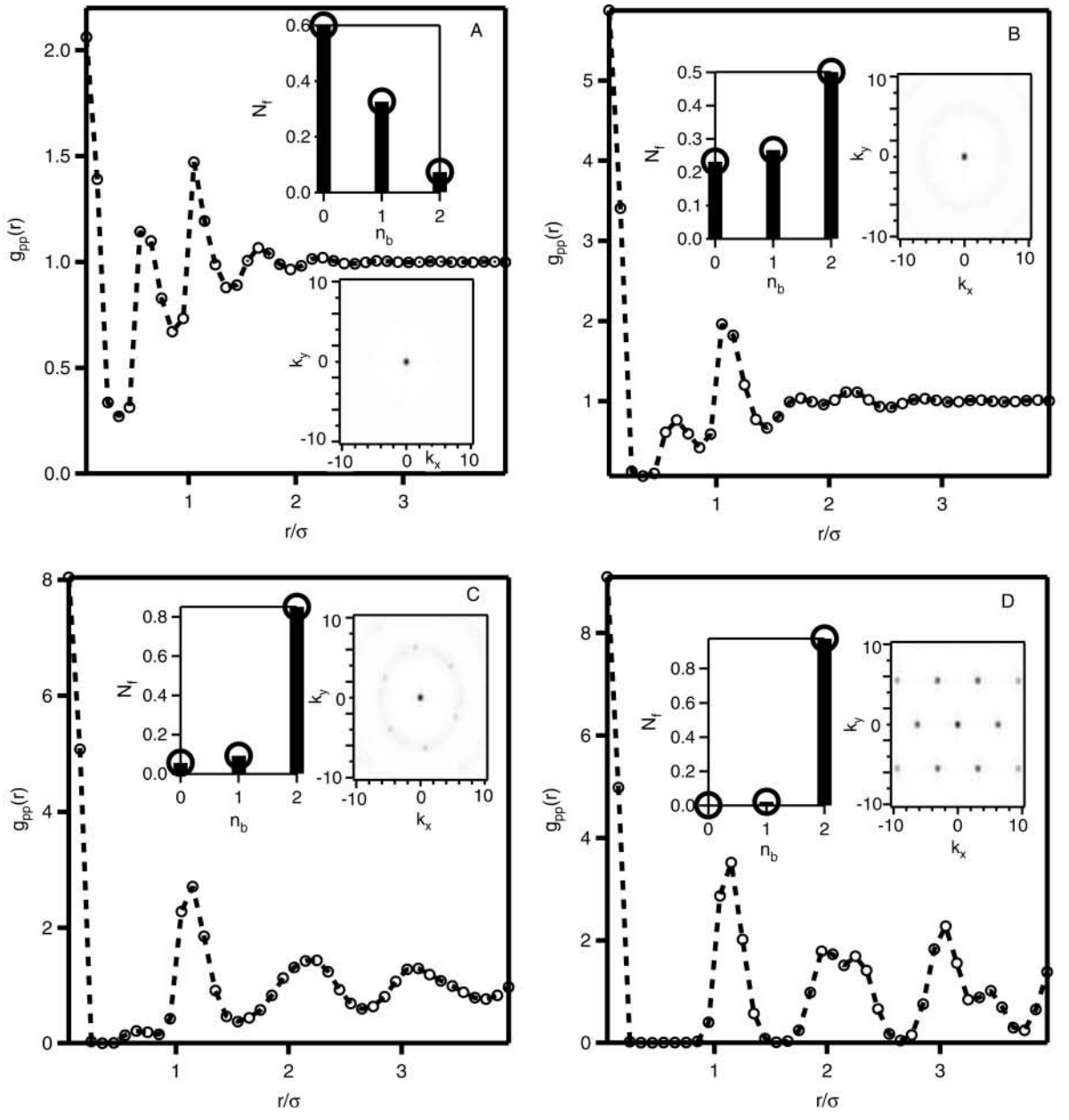

FIGURE 3 Structural properties of the systems with *p-p* interactions. The main figure shows the apex-apex pair correlation function, $g_{pp}(r)$, plotted against r/σ. One of the inset plots shows the fraction of the total number of sectors, N_f, plotted against the number of nearest sectors, n_b, and the other inset plot shows the contour plot of the structure factor, $S_{pp}(\mathbf{k})$, plotted against k_x and k_y. The four parts are for: (*A*) $\epsilon^*_{pp} = 1$ and $\phi = 0.63$; (*B*) $\epsilon^*_{pp} = 2.5$ and $\phi = 0.66$; (*C*) $\epsilon^*_{pp} = 4$ and $\phi = 0.64$; and (*D*) $\epsilon^*_{pp} = 4$ and $\phi = 0.68$.

$\phi = 0.68$ at the same value of $\epsilon^*_{pp} = 4$ (Fig. 3 *D*), $g_{pp}(r)$ shows sharp peaks even for large values of r, the structure factor has significant peaks at hexagonal positions, and the fraction of monomers that is part of trimers is almost equal to one.

The conclusions drawn from the structural features in Fig. 3 are confirmed from snapshots of the sectors. Fig. 4 depicts snapshots of final configurations corresponding to the cases depicted in Fig. 3. In Fig. 4 *A* we observe a disordered fluid of monomers, in *B* some but not all proteins are in trimeric form, in *C* most of the molecules are in trimeric form but a crystal is not clearly discernible, and in *D* the system is clearly a hexagonal lattice of trimers. Note that the system in part *C* is crystalline judging from the correlation functions, but this is not easily seen in the snapshot.

We conclude that a weak attraction between the *p-p* sites does not contribute to trimer formation even in very dense systems. As ϵ^*_{pp} is increased, it favors the formation of the trimers, and more so at higher area fractions. At higher values of the attractive well depth with $\epsilon^*_{pp} = 5$ and $\epsilon^*_{pp} =$

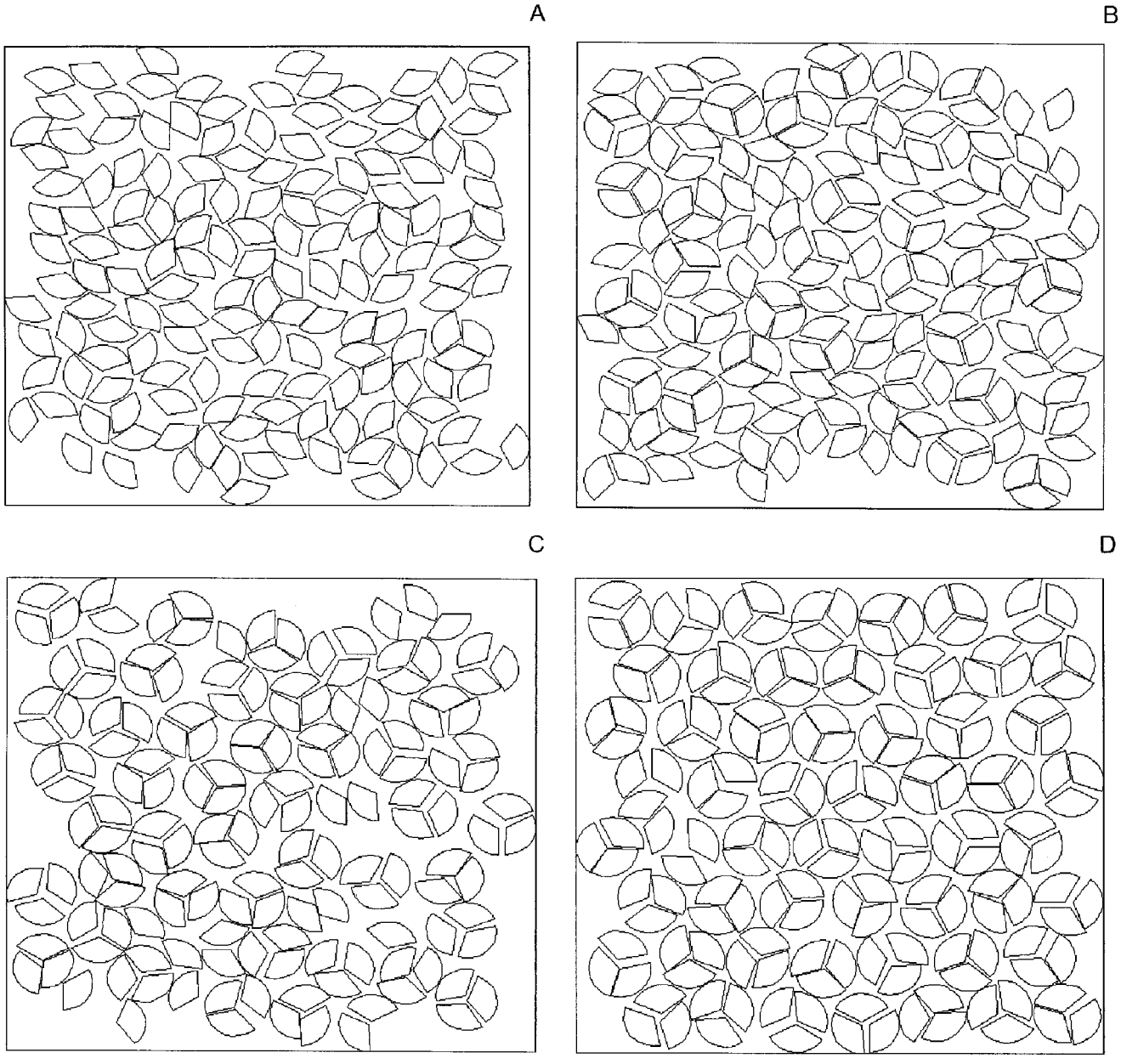

FIGURE 4 Snapshots of the final configurations for the four cases shown in Fig. 3, i.e., (*A*) $\epsilon^*_{pp} = 1$ and $\phi = 0.63$; (*B*) $\epsilon^*_{pp} = 2.5$ and $\phi = 0.66$; (*C*) $\epsilon^*_{pp} = 4$ and $\phi = 0.64$; and (*D*) $\epsilon^*_{pp} = 4$ and $\phi = 0.68$.

8, there are significant number of trimers present, even in the less dense systems. The formation of the hexagonal lattice begins at large values of ϕ, although it occurs at progressively lower values of ϕ as ϵ^*_{pp} is increased. Finally, a perfect hexagonal lattice is formed in the highly dense systems with strong enough attractions.

Fig. 5 depicts the fraction of trimers (fraction of sectors with $n_b = 2$), f, as a function of ϕ for different values of the attractive well depth, ϵ^*_{pp}. For $\epsilon^*_{pp} = 1$, the value of f reaches only 0.1 even at very high area fractions. As the strength of the attraction between the sites increases, trimers begin to form in the system. For $\epsilon^*_{pp} = 2.5$, ~50% of the molecules are in trimeric form for the highest concentration studied. For $\epsilon^*_{pp} = 4$ and $\epsilon^*_{pp} = 5$, f is initially low at smaller values of ϕ, but then reaches close to 100% (all trimers), at high values of ϕ. For $\epsilon^*_{pp} = 8$, the molecules exist predominantly as trimers for all area fractions.

The transition from the disordered monomeric state to the disordered trimeric state is smooth, but a characteristic area fraction for this transition can be located by considering the specific heat per sector, $C_v/\langle N\rangle k_B$. For $\epsilon^*_{pp} = 1$, the specific heat is a monotonically increasing function of ϕ. For $\epsilon^*_{pp} =$

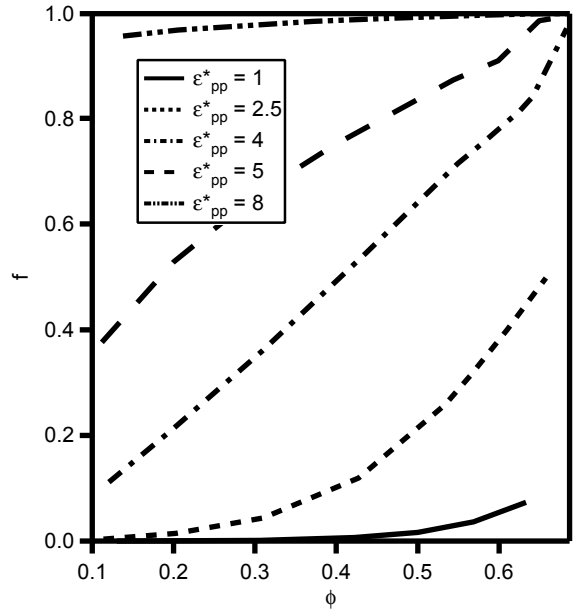

FIGURE 5 The fraction of trimers, *f,* plotted against the area fraction, ϕ, for different values of the attractive well depth, ϵ^*_{pp}, ranging from 1 to 8.

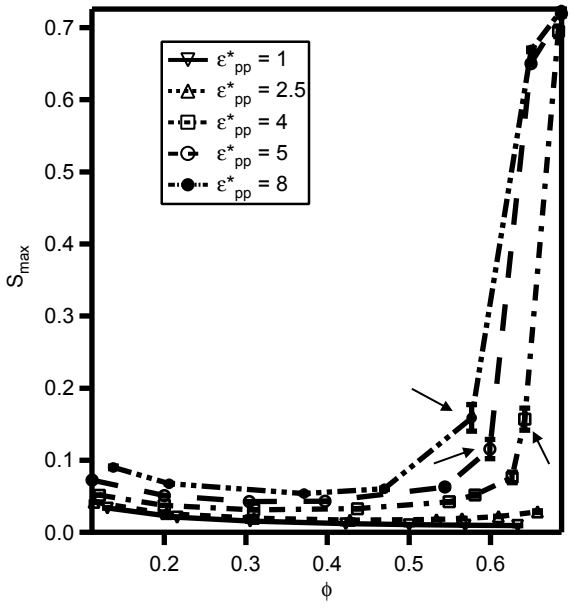

FIGURE 7 The parameter S_{max} plotted against the area fraction, ϕ, for different values of the attractive well depth, ϵ^*_{pp}, ranging from 1 to 8. The arrows indicate the transition points from the disordered trimer phase to the hexagonal lattice for $\epsilon^*_{pp} = 4$, $\epsilon^*_{pp} = 5$, and $\epsilon^*_{pp} = 8$.

2.5 and $\epsilon^*_{pp} = 4$, however, the specific heat shows a peak as a function of ϕ. This is shown in Fig. 6, which depicts the specific heat per sector as a function of ϕ, for $\epsilon^*_{pp} = 2.5$ and $\epsilon^*_{pp} = 4$. For these two cases, the peak occurs for $\phi = 0.54$ and $\phi = 0.3$, respectively. As the number of trimers in the system increases, the energy fluctuations of individual sectors in the trimeric units begin to decrease. Because the specific heat is related to the fluctuations in the energy, the

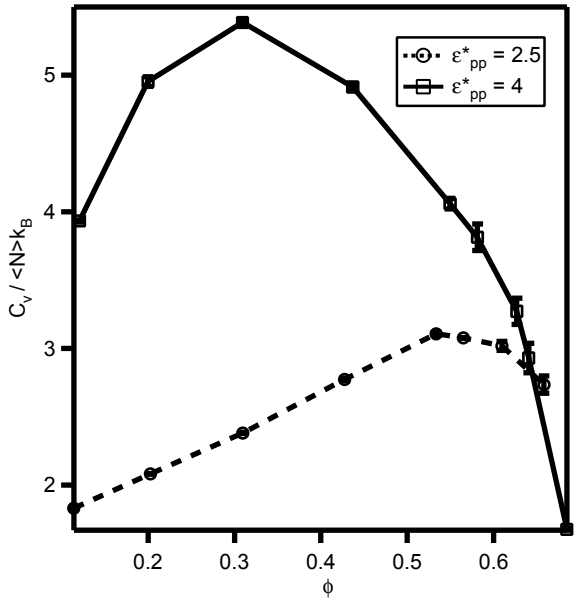

FIGURE 6 The specific heat per sector, $C_v/\langle N\rangle k_B$, plotted against the area fraction, ϕ, for $\epsilon^*_{pp} = 2.5$ and $\epsilon^*_{pp} = 4$.

ϕ at which there is a decrease in the specific heat can be attributed to the onset of a phase where the BR trimers are present in a significant number in the system. For $\epsilon^*_{pp} = 5$ and $\epsilon^*_{pp} = 8$ the specific heat decreases over the entire range of area fractions considered. This is due to the fact that at high values of ϵ^*_{pp}, the trimers are present in large numbers even in the most dilute systems studied.

The onset of formation of the hexagonal lattice can be quantitatively determined by considering the ratio of the intensity of the brightest spot in the structure factor to the intensity at the origin. We define a parameter S_{max}, which is the intensity of the most intense spot in $S_{pp}(\mathbf{k})$ with $\mathbf{k} \neq (0, 0)$ divided by the intensity of $S_{pp}(0, 0)$ (which is equal to $\langle N^2\rangle$). Fig. 7 is a plot of S_{max} as a function of ϕ for several values of ϵ^*_{pp}. For $\epsilon^*_{pp} = 1$ and 2.5, $S_{max} \approx 0$ for all values of ϕ, which indicates that there is no formation of a 2D hexagonal pattern for these values of ϵ^*_{pp}. For $\epsilon^*_{pp} = 4$, there is an increase in S_{max} at $\phi = 0.64$, which denotes the onset of formation of the crystal lattice. Similar increases in S_{max} are observed for $\epsilon^*_{pp} = 5$ and $\epsilon^*_{pp} = 8$, for $\phi = 0.6$ and $\phi = 0.57$, respectively. Because the increase in S_{max} appears to be sudden but small, we speculate that the liquid-solid transition in this system is weakly first-order. For the case of the hard disks, the transition from a liquid to the hexagonal crystal lattice is seen for $\phi = 0.56$ for our system sizes. This is consistent with the fact that for $\epsilon^*_{pp} = 8$, where the system has predominantly trimeric units over the entire range of concentrations, the transition to a crystal occurs for $\phi = 0.57$.

The above information can be summarized in a phase diagram with $1/\epsilon^*_{pp}$ plotted against the area fraction, ϕ (Fig.

8). There is one phase where the system is disordered with a significant number of monomers. Another phase consists of a significantly large number of BR trimers, but they are present in a random arrangement. Finally, there is the phase consisting of the 2D array of the BR timers in a hexagonal lattice. The value of ϕ at which the specific heat begins to decrease (Fig. 6) determines the transition from the disordered monomer phase to the disordered trimer phase. The parameter S_{max} (discussed in Fig. 7) determines the onset of the hexagonal lattice phase. At low values of ϵ^*_{pp} (high values of $1/\epsilon^*_{pp}$), the disordered phase of the monomers is present at all values of ϕ. As ϵ^*_{pp} increases, trimers begin to form even at low area fractions. The hexagonal assembly of the BR trimers forms at large values of ϵ^*_{pp} and in the highly dense systems, with the 2D crystal beginning to form at lower values of ϕ as ϵ^*_{pp} increases. The trimeric units in the

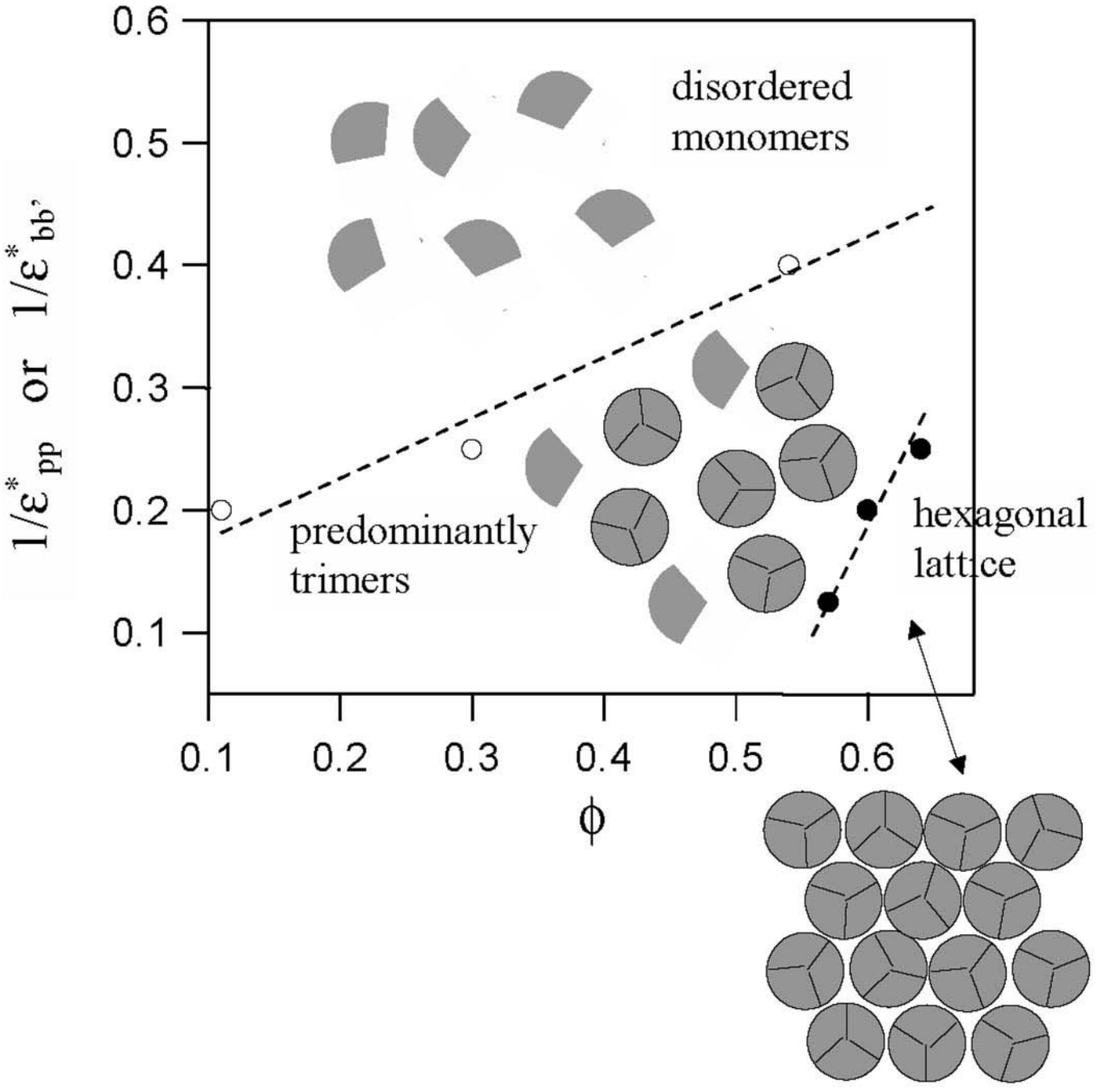

FIGURE 8 The phase diagram for the case of the *p-p* attractions with $1/\epsilon^*_{pp}$ plotted against the area fraction, ϕ. The three phases indicated are the disordered monomer phase, a predominantly trimer phase, and the hexagonal lattice assembly. The lines drawn through the points are merely to guide the eye.

crystal lattice are, however, not oriented in a specific direction and there is no rotational order in the hexagonal assembly as is observed in the PM.

Experimentally, it is much easier to modify side chains on the helices of the proteins than it is to modify the lipid molecules that bind to the proteins. Therefore, it is important to consider the effect of *b-b′* interactions instead of *p-p* interactions. Because each trimer can have a maximum of three *p-p* interactions and three *b-b′* interactions, the energy scales in these two cases are expected to be similar.

We find that the role of *b-b′* interactions in trimer formation and lattice assembly is almost identical to that of the *p-p* interactions. We perform a similar study, as described for *p-p* interactions, with $\epsilon^*_{bb'}$ ranging from 1 to 8, and ϕ ranging from 0.1 to 0.68. For example, consider the variation of the specific heat per sector, $C_v/\langle N\rangle k_B$, with ϕ for different values of the attractive well depth, $\epsilon^*_{bb'}$. For $\epsilon^*_{bb'} = 1$, the specific heat monotonically increases with ϕ. Fig. 9 depicts $C_v/\langle N\rangle k_B$ as a function of ϕ, for $\epsilon^*_{bb'} = 2.5$ and $\epsilon^*_{bb'} = 4$. As observed in the case of the *p-p* interactions, there is a peak in the specific heat. This peak occurs at values of $\phi = 0.54$ and $\phi = 0.3$, respectively, which are the same values for the area fractions at which peaks in the specific heat curves are seen in the case of *p-p* interactions. For $\epsilon^*_{bb'} = 5$ and $\epsilon^*_{bb'} = 8$, the specific heat decreases over the entire range of the area fractions considered, indicating that trimers are present even in less dense systems at high values of the attractive well depth.

The variation of the specific heat (for the case of the *b-b′* interactions) demonstrates that the transition from the disordered phase of the individual sectors, to the phase in which the trimers are predominantly present, occurs for the same values of ϕ as in the case of the *p-p* interactions. This suggests that the interactions between the internal lipids play a similar role in the formation of the BR trimers as do the protein-protein interactions. Because the formation of the hexagonal lattice (once the trimers are formed) is entropically driven in our model, the onset of the 2D crystal lattice can be expected to occur at similar values of ϕ for both *b-b′* and *p-p* interactions. This is confirmed from the variation of the parameter S_{max} with ϕ (not shown). It so transpires that the phase diagram in Fig. 8 for the *p-p* interactions holds for the *b-b′* interactions with ϵ^*_{pp} replaced by $\epsilon^*_{bb'}$! Note that, as in the case of *p-p* interactions, the crystal lattice is not orientationally ordered when $b - b'$ interactions are present.

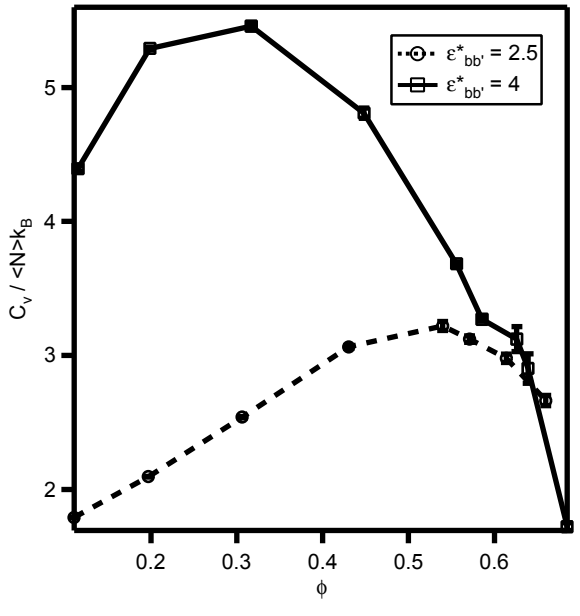

FIGURE 9 The specific heat per sector, $C_v/\langle N\rangle k_B$, plotted against the area fraction, ϕ, for $\epsilon^*_{bb'} = 2.5$ and $\epsilon^*_{bb'} = 4$.

The phase diagram depicted in Fig. 8 shows the importance of *p-p* or *b-b′* attractions, which is our model for interactions induced between the internal lipids or protein molecules, respectively, on the behavior of BR in membranes. There are two aspects that merit emphasis. The first is that the formation of trimers is energetically driven by the interactions between the monomers. The second is that the crystallization of trimers is entropically driven because in our model there is no attractive interaction whatsoever between trimers. There is one aspect of the PM, however, that the model does not reproduce. In the PM, the BR trimers also possess orientational order, i.e., the trimers all "point" in the same direction. This order is absent in the sector model *with only p-p or b-b′ interactions.*

To examine the impact of external lipids on the phase behavior, we now consider the role of *a-a′* interactions. For $\epsilon^*_{aa'} = 4$, $\epsilon^*_{pp} = \epsilon^*_{bb'} = 0$, and $\phi = 0.65$, the proteins are in predominantly monomeric form. Because the protein monomers form trimeric units for ϵ^*_{pp} or $\epsilon^*_{bb'} = 4$, $\epsilon^*_{aa'} = 0$, and $\phi = 0.65$, we conclude that interactions between the external lipids do not contribute to the formation of the BR trimers. Now consider the role of *a-a′* interactions in addition to *p-p* interactions. We do not determine the entire phase diagram for this set of interactions, but instead investigate select cases. The structural properties for $\epsilon^*_{pp} = 4$, $\epsilon^*_{aa'} = 1$, and $\phi = 0.68$ are depicted in Fig. 10. The $g_{pp}(r)$ is almost identical to that obtained for the case of only *p-p* interactions (Fig. 3 *D*), with $\epsilon^*_{pp} = 4$, $\epsilon^*_{aa'} = 0$, and $\phi = 0.68$. Similarly, almost all the BR molecules are found in the trimeric form, as is seen in the inset plot showing the fraction of the total number of sectors, N_f, plotted against $n_b = 0$, 1, and 2. However, incorporation of the *a-a′* interactions causes the trimers to prefer a specific orientation in the crystal lattice. Fig. 11 depicts the probability distribution function, $P(\theta)$, for the angle θ between the bisectors of any two sectors in the crystal, for $\epsilon^*_{pp} = 4$ and $\phi = 0.68$, with and without the *a-a′* interactions (i.e., $\epsilon^*_{aa'} = 1$ and $\epsilon^*_{aa'} = 0$, respectively). Note that since the difference in the angle θ between the bisectors of any two sectors in the same trimeric unit is always 120° (because the sectors do not overlap with each other), this "self" contribution is not

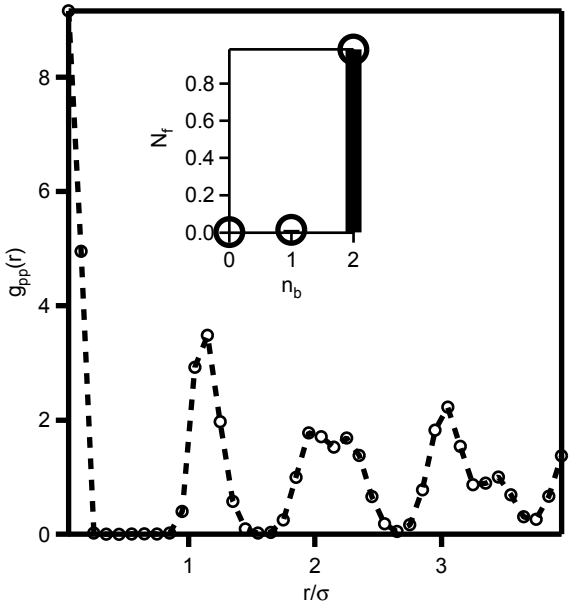

FIGURE 10 Structural properties of the system for $\epsilon^*_{aa'} = 1$, $\epsilon^*_{pp} = 4$, and $\phi = 0.68$. The main figure depicts the radial distribution function, $g_{pp}(r)$, plotted against r/σ. The inset plot shows the fraction of the total number of sectors, N_f, plotted against the number of nearest sectors, n_b. Note the similarity between this figure and Fig. 3 *D*.

included in $P(\theta)$. Without the *a*-*a*′ interactions, $P(\theta)$ is nearly independent of θ and is $\simeq$0.0056 (which is equal to 1/180°). There is a slight oscillation in $P(\theta)$ in this case which is because of the 120° form of the BR sectors.

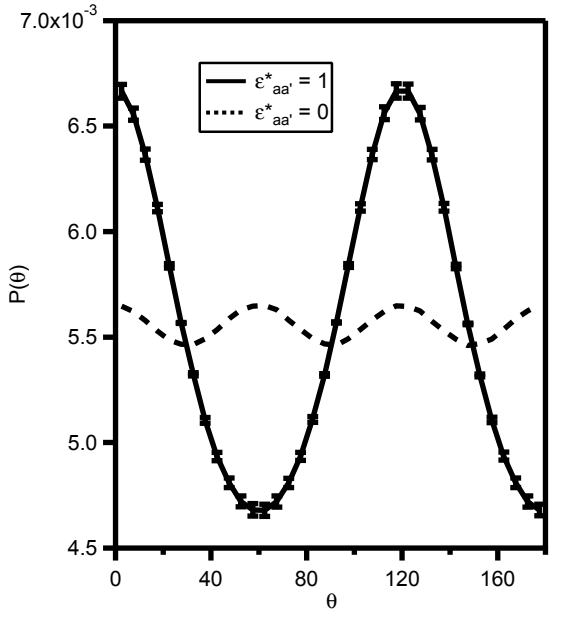

FIGURE 11 Probability distribution, $P(\theta)$, for the angle, θ, between the bisectors of any two sectors for $\epsilon^*_{pp} = 4$ and $\phi = 0.68$ with $\epsilon^*_{aa'} = 0$ and 1. Note that the *a*-*a*′ interaction causes the trimers to orient in a specific direction.

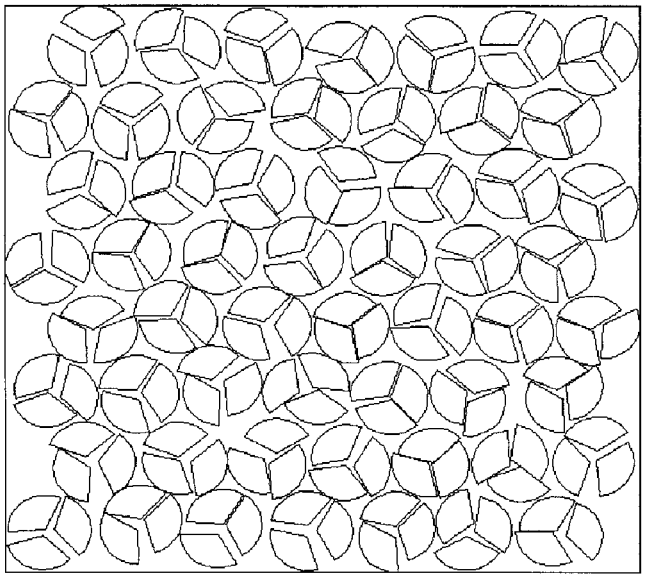

FIGURE 12 Snapshot of the final configuration of the system for $\epsilon^*_{aa'} = 1$, $\epsilon^*_{pp} = 4$, and $\phi = 0.68$ showing the hexagonal order and the specific orientation of the trimers.

However, in the presence of the *a*-*a*′ interactions, there is a specific orientational order in the crystal. This is clear from the fact that $P(\theta)$ is peaked at 0° and 120°. This is the case when the trimeric units are oriented in a specific direction in the crystal, and the difference in the angle θ between any two sectors (belonging to different trimers) is nearly either 0° or 120°. (The errorbars for $\epsilon^*_{aa'} = 0$ are not shown because they are much smaller than the data points). These conclusions are confirmed in Fig. 12, which depicts a snapshot of the final configuration of the crystal lattice with the *a*-*a*′ interactions present. The structure is very similar to that seen without the *a*-*a*′ interactions (Fig. 4 *D*), except that there is clearly a preferred orientation for the BR trimers.

An important parameter in our model is the range of attractive interactions, λ, and we present some preliminary calculations to investigate the effect of this parameter. To this end, we perform simulations for $\phi > 0.65$ and for $\lambda = 0.1$, 0.3, and 0.4. For $\lambda = 0.2$, the system is crystalline for these area fractions and ϵ^*_{pp} or $\epsilon^*_{bb'} > 4$. Note that if $\lambda \geq 0.5$, the sectors do not experience an orientation-dependent attraction, and we expect ordered states to be less favorable. For the case of *p*-*p* interactions, the qualitative properties appear to be relatively insensitive to the value of the range of interaction. The system still undergoes transitions from monomeric to trimeric to crystalline states, although this occurs at high values of ϵ^*_{pp} as λ is decreased. For the case of *b*-*b*′ interactions, however, we find that increasing the range of interactions can destroy the ordered state. For $\lambda = 0.1$ and 0.2 we observe ordered states, although a higer value of $\epsilon^*_{bb'}$ is required for $\lambda = 0.1$. For $\lambda = 0.3$ and 0.4, however, we do not observe ordered states. This indicates

TABLE 1 When $\sigma/\sqrt{2} \le R < \sigma$

Range of α	β_1	β_2
$\alpha < -\pi/3 - \theta$	no overlap	no overlap
$-\pi/3 - \theta \le \alpha < -\pi/3$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$	$\pi + \theta$
$-\pi/3 \le \alpha < -\pi/3 + \theta$	$\pi + ANG_{b1}(R,\alpha + \pi/3)$	$\pi + \theta$
$-\pi/3 + \theta \le \alpha < \pi/3 - \theta$	$\pi - \theta$	$\pi + \theta$
$\pi/3 - \theta \le \alpha < \pi/3$	$\pi - \theta$	$\pi + ANG_{b1}(R, \alpha - \pi/3)$
$\pi/3 \le \alpha < \pi/3 + \theta$	$\pi - \theta$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$
$\pi/3 + \theta \le \alpha$	no overlap	no overlap

TABLE 2 When $\sigma/2 \le R < \sigma/\sqrt{2}$

Range of α	β_1	β_2
$\alpha < -\pi/3 - \theta_1$	no overlap	no overlap
$-\pi/3 - \theta_1 \le \alpha < -\pi/3 - \theta$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$	$\pi + ANG_{a2}(R, \alpha + \pi/3)$
$-\pi/3 - \theta \le \alpha < -\pi/3 - \theta_3$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$	$\pi + ANG_{b1}(R, \alpha + \pi/3)$
$-\pi/3 - \theta_3 \le \alpha < -\pi/3$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$	$\pi + \theta_1$
$-\pi/3 \le \alpha < -\pi/3 + \theta_3$	$\pi + ANG_{b1}(R, \alpha + \pi/3)$	$\pi + \theta_1$
$-\pi/3 + \theta_3 \le \alpha < \pi/3 - \theta_3$	$\pi - \theta_1$	$\pi + \theta_1$
$\pi/3 - \theta_3 \le \alpha < \pi/3$	$\pi - \theta_1$	$\pi + ANG_{b1}(R, \alpha - \pi/3)$
$\pi/3 \le \alpha < \pi/3 + \theta_3$	$\pi - \theta_1$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$
$\pi/3 + \theta_3 \le \alpha < \pi/3 + \theta$	$\pi + ANG_{b1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$
$\pi/3 + \theta \le \alpha < \pi/3 + \theta_1$	$\pi + ANG_{a2}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$
$\pi/3 + \theta_1 \le \alpha$	no overlap	no overlap

TABLE 3 When $\sigma/4 \le R < \sigma/2$

Range of α	β_1	β_2
$\alpha < -5\pi/6$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\alpha = -5\pi/6$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + \theta_c$
$-5\pi/6 < \alpha < -\pi/3 - \theta$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{a2}(R\ \alpha + \pi/3)$
$-\pi/3 - \theta \le \alpha < -2\pi/3$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{b1}(R, \alpha + \pi/3)$
$-2\pi/3 \le \alpha < -\pi/3 - \theta_c$	π	$\pi + ANG_{b1}(R, \alpha + \pi/3)$
$-\pi/3 - \theta_c \le \alpha < -\pi/3$	π	$2\pi + ANG_{b1}(R, \alpha + \pi/3)$
$-\pi/3 \le \alpha < \pi/3$	always overlap	always overlap
$\pi/3 \le \alpha < \pi/3 + \theta_c$	$ANG_{b1}(R, \alpha - \pi/3)$	π
$\pi/3 + \theta_c \le \alpha < 2\pi/3$	$\pi + ANG_{b1}(R, \alpha - \pi/3)$	π
$2\pi/3 \le \alpha < \pi/3 + \theta$	$\pi + ANG_{b1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\pi/3 + \theta \le \alpha < 5\pi/6$	$\pi + ANG_{a2}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\alpha = 5\pi/6$	$\pi - \theta_c$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\alpha > 5\pi/6$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$

TABLE 4 When $R < \sigma/4$

Range of α	β_1	β_2
$\alpha < -5\pi/6$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\alpha = -5\pi/6$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + \theta_c$
$-5\pi/6 < \alpha < -\pi/3 - 0$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{a2}(R, \alpha + \pi/3)$
$-\pi/3 - \theta < \alpha < -\pi/3 - \theta_c$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{b1}(R, \alpha + \pi/3)$
$-\pi/3 - \theta_c \le \alpha < -2\pi/3$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$2\pi + ANG_{b1}(R, \alpha + \pi/3)$
$-2\pi/3 \le \alpha < -\pi/3$	π	$2\pi + ANG_{b1}(R, \alpha + \pi/3)$
$-\pi/3 \le \alpha < \pi/3$	always overlap	always overlap
$\pi/3 < \alpha < 2\pi/3$	$ANG_{b1}(R, \alpha - \pi/3)$	π
$2\pi/3 \le \alpha < \pi/3 + \theta_c$	$ANG_{b1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\pi/3 + \theta_c \le \alpha < \pi/3 + \theta$	$\pi + ANG_{b1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\pi/3 + \theta \le \alpha < 5\pi/6$	$\pi + ANG_{a2}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\alpha = 5\pi/6$	$\pi - \theta_c$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$
$\alpha > 5\pi/6$	$\pi + ANG_{a1}(R, \alpha - \pi/3)$	$\pi + ANG_{a1}(R, \alpha + \pi/3)$

that the protein-protein interactions have to be short-ranged, while the internal lipid-lipid interactions can be slightly long-ranged to aid the formation of trimers.

SUMMARY AND CONCLUSIONS

We study the self-assembly of the bacteriorhodopsin (BR) monomers in the purple membrane (PM) using grand canonical Monte Carlo simulations of a simple 2D model of protein and lipids. The protein molecules are modeled as 120° sectors of a circle. In addition to excluded volume interactions, which is enforced by requiring that two sectors cannot overlap, we include site-site square-well attractions to mimic protein-protein interactions and lipid-mediated interactions. Three types of site-site attractions are considered: attractions between the apexes of sectors are meant to model interactions between lipids bound to the proteins and present in the interior of the BR trimer, attractions between the midpoints of the sides of the sectors are meant to model interactions between adjacent α-helices in the trimer, and interactions between sites on the circumference of the sectors are meant to mimic lipid-mediated interactions between α-helices on the exterior of the protein.

We find that in the absence of attractive interactions, or if the attractive interactions are weak, the proteins are in the monomeric disordered state for all concentrations. If a strong enough attraction between the apexes of the sectors (internal lipid-lipid interaction) is incorporated, there is a smooth transition from a disordered monomeric liquid to a fluid of trimers as the protein concentration is increased. This is manifested by a peak in the specific heat, in addition to smooth changes in the fraction of proteins that are part of a cluster of three proteins. As the protein concentration is increased further, there is a sharp transition from a disordered fluid of trimeric units to a hexagonally ordered crystal of trimers. We speculate that this transition is weakly first-order. As the strength of the attraction is increased, these two phase transitions occur at progressively lower protein concentrations. The incorporation of protein-protein interactions (between midpoints of the two sides of the sectors) instead of internal lipid-lipid interactions leaves the phase diagram unchanged. In fact, the concentrations at which the monomer to disordered trimer and disordered trimer to crystal transitions occur are almost identical whether protein-protein interactions or internal lipid-lipid interactions are present. In both these cases, the resulting crystal does not possess orientational order, the trimers are randomly oriented in space. This suggests that although protein-protein and internal lipid-lipid interactions are important factors for the formation of the trimeric units and the crystal, additional interactions are present in the PM. The incorporation of lipid-mediated interactions between two external helices on the proteins completes the story. These interactions do not affect the transition from monomers to trimers or the transition from disordered trimers to the crystal. However, when the crystal is formed these interactions cause the trimers to prefer a specific orientation, as is observed in the PM.

The present work thus elucidates the role of various interactions in the assembly of BR in the PM. We conclude that protein-protein and lipid-lipid interactions of the internal lipids contribute equally to the formation of the trimeric BR units and the 2D hexagonal arrangement in the PM. This is of possible importance in the crystallization of membrane proteins whose structure is not known. One can imagine constructing a bilayer of lipid molecules with proteins and then attempting to promote crystallization by changing the nature of specific lipid molecules. We also predict that trimer formation is a necessary precursor to crystallization of BR in the PM, i.e., the monomer and crystal do not coexist. Although this is not altogether unexpected, there has been some discussion in the literature regarding this issue (Krebs and Isenbarger, 2000), and this work provides theoretical evidence in this regard. Our model also suggests that the role of the external lipids in the PM is to orient the BR trimers in a specific direction. While we are not familiar with physiological reasons why the BR trimers should be oriented thus, this is a curious fact that merits study. Supramolecular assembly of optically active materials into thin film architecture that possesses improved optical functionality compared to the bulk state (He et al., 1999), is of significant practical importance. The photoelectric conversion efficiency of a BR film is largely dependent on the degree of orientation of BR in the film.

We should point out that although our model is extremely simple, its success in describing the properties of the PM suggests that a phenomenological approach of this type can provide insight and complement experimental findings in a useful fashion. There are several other examples of naturally occurring membrane protein complexes that self-assemble into crystals (Koepke et al., 1996; Prince et al., 1997) and such models could be used in these cases to understand the interactions that provide the driving force for the observed structure.

In conclusion, we have presented a minimal model that displays the essential features of self-assembly of bacteriorhodopsin in the purple membrane. The model has only two parameters which, in principle, could be determined if experimental results for the phase behavior become available. We hope our work will motivate such studies. In addition, the model can now be used to study other aspects of self-assembly such as the effect of external fields and the kinetics of crystal formation.

APPENDIX

This section outlines the conditions when two sectors overlap. For a given separation between the apexes of two sectors, the sectors overlap for certain relative angles of the two sectors (Fig. 2). The detailed procedure for the check of the overlap of sectors is described below.

We align two sectors of same size such that the apex of the sectors is located at $(-R/2, 0)$ (referred to as the "left sector") and $(R/2, 0)$ (referred to as the "right sector"), respectively, where R is the distance between the apexes of the two sectors. Two circles, with centers coinciding with the apexes and radii equal to $\sigma/2$, are drawn as shown in Fig. 2. The sectors are shown by shaded regions in the respective circles. The angles α ($-\pi \le \alpha < \pi$) and θ ($0 < \theta < \pi/2$) are defined as the angle between the bisector of the left sector and the horizontal and the angle between the line OC and the horizontal, respectively. (Note that $R = \sigma \cos\theta$.)

The main procedure is outlined as follows. The intersection of the left sector with the intersection area of the two circles is determined first. Then the smallest sector with its apex located at the center of the right circle and spanning the intersection area determined in the previous step is obtained (note that the radius of this sector is equal to $\sigma/2$). The smallest sector is described by two angles, β_1 and β_2 ($0 \le \beta_1 \le \beta_2 < 2\pi$), where β_1 and β_2 are the angles formed by the two sides of the smallest sector with the horizontal. Finally, the overlap between the smallest sector and the right sector is determined. To obtain the β values, the configurations are divided into five cases depending on R and each case is subdivided depending on α. The β values for each case are described below in detail.

When $R \ge \sigma$, there is no overlap between the two sectors irrespective of their orientation. When $\sigma/\sqrt{2} \le R < \sigma$, $\sigma/2 \le R. < \sigma/\sqrt{2}$, $\sigma/4 \le R < \sigma/2$ and $R < \sigma/4$, the various subcases are enumerated in Table 1–4. The angles θ_1 and θ_3 are defined as shown in Fig. 2, and θ_c is defined as $\cos^{-1}(2\cos\theta)$. The functions ANG_{a1}, ANG_{a2}, and ANG_{b1} are defined as follows:

$$ANG_{a1}(R, \theta) = \tan^{-1}\left(\tan\theta \frac{\sqrt{\frac{1+\tan^2\theta}{4} - R^2\tan^2\theta} - R}{\sqrt{\frac{1+\tan^2\theta}{4} - R^2\tan^2\theta} + R\ \tan^2\theta}\right) \quad (6)$$

$$ANG_{a2}(R, \theta) = \tan^{-1}\left(\tan\theta \frac{\sqrt{\frac{1+\tan^2\theta}{4} - R^2\tan^2\theta} + R}{\sqrt{\frac{1+\tan^2\theta}{4} - R^2\tan^2\theta} - R\tan^2\theta}\right) \quad (7)$$

$$ANG_{b1}(R, \theta) = \tan^{-1}\left(\frac{\sin\theta}{\cos\theta - 2R}\right) \quad (8)$$

We thank Professor Mark Krebs for introducing us to this problem, for several useful discussions, and for critical comments regarding the manuscript. We also thank Professor Jim Weisshaar and Mr. Christoph Weise for discussions, and Professor Ken Dill for an interesting talk that inspired the two-dimensional model used in this work.

We gratefully acknowledge support from the National Science Foundation through Grant CHE 9732604.

REFERENCES

Alder, B. J., and T. E. Wainwright. 1962. Phase transitions in elastic disks. *Phys. Rev.* 127:359–361.

Allen, M. P., and D. J. Tildesley, editors. 1987. Computer Simulation of Liquids. Clarendon Press, Oxford.

Baudry, J., E. Tajkhorshid, F. Molnar, J. Phillips, and K. Schulten. 2001. Molecular dynamics study of bacteriorhodopsin and the purple membrane. *J. Phys. Chem. B.* 105:905–918.

Belrhali, H., P. Nollert, A. Royant, C. Menzel, J. Rosenbusch, E. M. Landau, and E. Pebay-Peyroula. 1999. Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at 1.9 angstrom resolution. *Structure.* 7:909–917.

Booth, P. J. 2000. Unraveling the folding of bacteriorhodopsin. *Biochim. Biophys. Acta.* 1460:4–14.

Edholm, O., O. Berger, and F. Jähnig. 1995. Structure and fluctuations of bacteriorhodopsin in the purple membrane: a molecular dynamics study. *J. Mol. Biol.* 250:94–111.

Essen, L. O., R. Siegert, W. D. Lehmann, and D. Oesterhelt. 1998. Lipid patches in membrane protein oligomers: crystal structure of the bacteriorhodopsin-lipid complex. *Proc. Natl. Acad. Sci. USA.* 95: 11673–11678.

Frenkel, D., and B. Smit, editors. 1996. Understanding Molecular Simulations: From Algorithms to Applications. Academic Press, San Diego.

Fyfe, P. K., K. E. McAuley, A. W. Roszak, N. W. Isaacs, R. J. Cogdell, and M. R. Jones. 2001. Probing the interface between membrane proteins and membrane lipids by x-ray crystallography. *Trends Biochem. Sci.* 26:106–112.

Grogorieff, N., T. A. Ceska, K. H. Downing, J. M. Baldwin, and R. Henderson. 1996. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* 259:393–421.

Halperin, B. I., and D. R. Nelson. 1978. Theory of two-dimensional melting. *Phys. Rev. Lett.* 41:121–124.

He, J.-A., L. Samuelson, L. Li, J. Kumar, and S. K. Tripathy. 1999. Bacteriorhodopsin thin film assemblies: immobilization, properties and applications. *Advanced Materials.* 11:435–446.

Henderson, R., J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing. 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* 213: 899–929.

Henderson, R., and P. N. T. Unwin. 1975. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature.* 257:28–32.

Isenbarger, T. A., and M. P. Krebs. 1999. Role of helix-helix interactions in assembly of the bacteriorhodopsin lattice. *Biochemistry.* 38: 9023–9030.

Jaster, A. 1999. Computer simulations of the two-dimensional melting transition using hard disks. *Phys. Rev. E.* 59:2594–2602.

Kimura, Y., D. G. Vassylyev, A. Miyazawa, A. Kidera, M. Matsushima, K. Mitsuoka, K. Murata, T. Hirai, and Y. Fujiyoshi. 1997. Surface of bacteriorhodopsin revealed by high-resolution electron crystallography. *Nature.* 389:206–211.

Koepke, J., X. C. Hu, C. Muenke, K. Schulten, and H. Michel. 1996. The crystal structure of the light-harvesting complex II (B800–850) from Rhodospirillum molischianum. Structure. 4:581–597.

Koltover, I., J. O. Raedler, T. Salditt, K. J. Rothschild, and C. R. Safinya. 1999. Phase behavior and interactions of the membrane-protein bacteriorhodopsin. *Phys. Rev. Lett.* 82:3184–3187.

Kosterlitz, J. M., and D. J. Thouless. 1973. Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. Chem. C.* 6:1181–1203.

Krebs, M. P., and T. A. Isenbarger. 2000. Structural determinants of purple membrane assembly. *Biochim. Biophys. Acta.* 1460:15–26.

Krebs, M. P., W. Li, and T. P. Halembeck. 1997. Intermembrane substitutions in helix D of bacteriorhodopsin disrupt the purple membrane. *J. Mol. Biol.* 267:172–183.

Lanyi, J. K. 2000. Bacteriorhodopsin. *Biochim. Biophys. Acta.* 1460:1–3.

Lanyi, J. K., and H. Luecke. 2001. Bacteriorhodopsin. *Curr. Opin. Struct. Biol.* 11:415–419.

Luecke, H., B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi. 1999. Structure of bacteriorhodopsin at 1.55 angstrom resolution. *J. Mol. Biol.* 291:899–911.

McQuarrie, D. A., editor. 1976. Statistical Mechanics. Harper & Row, New York.

Mermin, N. D., and H. Wagner. 1966. Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models. *Phys. Rev. Lett.* 17:1133–1136.

Oesterhelt, D., C. Bräuchle, and N. Hampp. 1991. Bacteriorhodopsin: a biological material for information processing. *Q. Rev. Biophys.* 24: 425–478.

Osterhelt, D., and W. Stoeckenius. 1973. Functions of a new photoreceptor membrane. *Proc. Natl. Acad. Sci. USA.* 70:2853–2857.

Pebay-Peyroula, E., and J. P. Rosenbusch. 2001. High-resolution structures and dynamics of membrane protein-lipid complexes: a critique. *Curr. Opin. Struct. Biol.* 11:427–432.

Prince, S. M., M. Z. Papiz, A. A. Freer, G. McDermott, A. M. Hawthornthwaite-Lawless, R. J. Cogdell, and N. W. Isaacs. 1997. Apoprotein structure in the LH2 complex from *Rhodopseudomonas acidophilla* strain 10050: modular assembly and protein pigment interactions. *J. Mol. Biol.* 268:412–423.

Sabra, M. C., J. C. M. Uitdehaag, and A. Watts. 1998. General model for lipid-mediated two-dimensional array formation of membrane proteins: application to bacteriorhodopsin. *Biophys. J.* 75:1180–1188.

Sternberg, B., A. Watts, and Z. Cejka. 1993. Lipid-induced modulation of the protein packing in 2-dimensional crystals of bacteriorhodopsin. *J. Struct. Biol.* 110:196–204.

Watts, A. 1995. Bacteriorhodopsin: the mechanism of 2D-array formation and the structure of the retinal in the protein. *Biophys. Chem.* 55: 137–151.

Weik, M., H. Patzelt, G. Zaccai, and D. Oesterhelt. 1998. Localization of glycolipids in membranes by in vivo labeling and neutron diffraction. *Mol. Cell.* 1:411–419.

Young, A. P. 1979. Melting and the vector Coulomb gas in two dimensions. *Phys. Rev. B.* 19:1855–1866.