Analysis of Simulated NMR Order Parameters for Lipid Bilayer Structure Determination

Horia I. Petrache,* Kechuan Tu,# and John F. Nagle*§

*Department of Physics and §Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, and #Department of Pharmacological Chemistry, University of California, San Francisco, California 94143 USA

ABSTRACT The conventional formula for relating CD_2 average order parameters $\langle S_n \rangle$ to average methylenic travel $\langle D_n \rangle$ is flawed when compared to molecular dynamics simulations of dipalmitoylphosphatidylcholine. Inspired by the simulated probability distribution functions, a new formula is derived that satisfactorily relates these quantities. This formula is used to obtain the average chain length $\langle L_C \rangle$, and the result agrees with the direct simulation result for $\langle L_C \rangle$. The simulation also yields a hydrocarbon thickness $2\langle D_C \rangle$. The result $\langle L_C \rangle = \langle D_C \rangle$ is consistent with a model of chain packing with both early chain termination and partial interdigitation of chains from opposing monolayers. The actual simulated area per lipid $\langle A \rangle$ is easily obtained from the order parameters. However, when this method is applied to NMR order parameter data from dimyristoylphosphatidylcholine, the resulting $\langle A \rangle$ is 10% larger than the currently accepted value.

INTRODUCTION

Simple lipid bilayers have been a challenge for quantitative structure determination. For example, experimental values for the average area per lipid $\langle A \rangle$ for fully hydrated dipalmitoylphosphatidylcholine (DPPC) at 50°C have ranged from 56 to 72 Å^2 (Nagle, 1993), with corresponding uncertainty in the average thickness of the hydrocarbon core, $2\langle D_C \rangle$. The same range of uncertainty has come from the two primary experimental methods, namely x-ray diffraction and NMR. Recently, primary emphasis has been on the x-ray technique, where $\langle A_{DPPC} \rangle = 62.9 \pm 1.3$ Å^2 was obtained (Nagle et al., 1996). This result was in agreement with a reanalysis of NMR order parameter data, which gave $\langle A_{DPPC} \rangle = 62 \pm 2$ Å^2 (Nagle, 1993). However, it was emphasized in that reanalysis that there were assumptions that could and should be tested with simulations. This paper carries forward that program.

Simulations give a much more detailed view than any experiments on lipid bilayers (Tobias et al., 1997; Tieleman et al., 1997). However, given the uncertainties in force fields and the restriction to nanosecond time scales, one should not necessarily expect simulations to obtain accurate values for all quantities of interest. The use of simulations that we envision is to test relations between simulated microscopic quantities that cannot be obtained from experiment. Such relations are commonly used to obtain quantities of interest from raw data. Simulations can also inspire new relations, as we show in this paper. This use of simulations does not require that all of the force fields be exactly correct or that the simulated quantities of experimental interest agree perfectly with experiment. The interplay of simulations that is envisioned is that simulations will help guide experimental analysis, which will then provide more accurate quantities of interest, which will then help tune the force fields used in the simulations.

The main relation studied in this paper is between the average order parameters $\langle S_n \rangle$ and the average travel per methylene $\langle D_n \rangle$ along the normal to the bilayer. The conventional relation (Schindler and Seelig, 1975; Thurmond et al., 1991; Nagle, 1993) can essentially be written as

$$\langle D_n \rangle / D_M = (1 - 2\langle S_n \rangle)/2, \quad (1)$$

where D_M is the maximum possible travel. A different relation has also been used (De Young and Dill, 1988):

$$\langle D_n \rangle / D_M = (1 - 4\langle S_n \rangle)/3. \quad (2)$$

However, a recent simulation (Berger et al., 1997) shows (their figure 7) that neither relation is especially good, and this paper reports essentially the same result for a different simulation. Even though the values of $\langle A \rangle$ that were obtained in their simulation using the method of Nagle (1993) were in fairly good agreement with the actual simulated $\langle A \rangle$, it is important to elucidate the errors, including fortuitous compensations.

The simulation we analyze in this paper has been described previously (Tu et al., 1995). Briefly, the simulation was performed at constant number, pressure, and temperature (NPT), with $N = 32$ DPPC lipids per monolayer and $n_W = 28$ waters/lipid, $P = 0$, and $T = 50$°C, using a Nose–Hoover thermostat. The SPC/E water potential and an all-atom description of the lipids were used in the CHARMM program, with Ewald sums for long-range interactions. Since publication (Tu et al., 1995) the simulation has been continued to 2 ns.

METHYLENE TRAVEL

The first main result is illustrated in Fig. 1, which plots the conventional CD_2 order parameters S_n versus the travel of

Received for publication 22 October 1998 and in final form 5 February 1999.

Address reprint requests to Dr. John F. Nagle, Department of Physics, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213. Tel.: 412-268-2764; Fax: 412-681-0648; E-mail: naglet@andrew.cmu.edu.

0006-3495/99/05/2479/09 $2.00

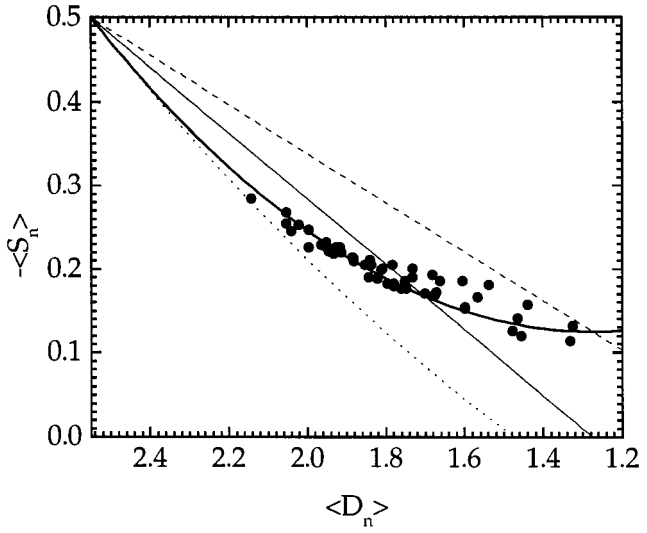

FIGURE 1 Symbols: simulated average order parameters $\langle S_n \rangle$ versus average molecular travel $\langle D_n \rangle$ (in Å). Curved solid line: New Eq. 24. Straight solid line: conventional Eq. 1. Dashed line: Eq. 2. Dotted line: Eq. 14.

the *n*th methylene along the bilayer normal. Each simulated data point in this figure gives molecular dynamical averages for one specific methylene, with the different points corresponding to different carbon numbers *n*, different chains (sn-1 and sn-2), and different monolayers (upper and lower) in the simulated bilayer. The solid straight line shows that the conventional prediction (Nagle, 1993) works fairly well for values of the order parameter $\langle S_n \rangle \approx -0.2$ that are close to the experimental plateau region, but the slope of the data is clearly smaller than the conventional slope. This point has previously been made for a different simulation by Berger et al. (1997); they presented essentially this type of figure (except that they plotted the molecular order parameter instead of $\langle S_n \rangle$) with very similar results. Our first main result is the derivation of a new formula that gives the result shown by the solid curve in Fig. 1.

The derivation of the new formula was inspired by detailed results of the simulation that have not previously been presented. As with any fluctuating statistical mechanical ensemble, one should consider distribution functions. A fairly general one that we have analyzed for the hydrocarbon chains is denoted $p_n(z, D)$. The discrete index *n* is the carbon number, which goes from 2 for the methylene next to the carbonyl to 16 for the terminal methyl. The continuous variable *z* is the distance along the bilayer normal, where $z = 0$ is the center of the bilayer and with the positive sign for the direction toward the headgroup region for that chain. The instantaneous position of carbon *n* is denoted by z_n. The continuous variable *D* is the methylenic travel; for carbon number *n* the instantaneous travel D_n is defined by

$$D_n = z_{n-1} - z_{n+1}, \tag{3}$$

as illustrated in Fig. 2. From the distribution function

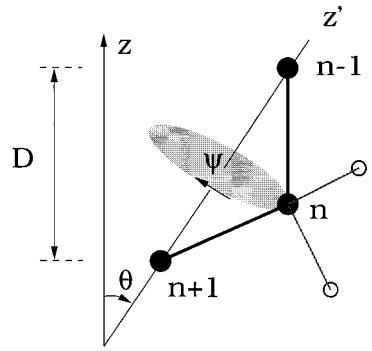

FIGURE 2 For the *n*th CH_2 group, *D* defines the travel along the bilayer normal *z*, θ defines the local methylenic tilt, and ψ defines the rotation about the local methylene axis *z*′. Note that the local axis is generally different for each carbon *n*.

$p_n(z, D)$ the average position of carbon *n* is given as

$$\langle z_n \rangle = \iint z p_n(z, D) \mathrm{d}z \, \mathrm{d}D. \tag{4}$$

Fig. 3 shows that $\langle z_n \rangle$ has upward curvature. This corresponds to D_n decreasing with *n*, which conforms to the expectation that the chains become more disordered near the terminal methyl end.

For many purposes it suffices to consider only the reduced distribution function

$$p_n(D) = \int p_n(z, D) \mathrm{d}z. \tag{5}$$

For example, the average travel of carbon *n*, irrespective of where it is, is then given by

$$\langle D_n \rangle = \int D p_n(D) \mathrm{d}D. \tag{6}$$

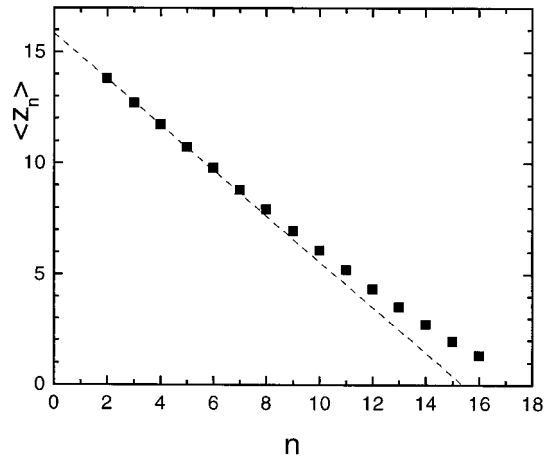

FIGURE 3 Average carbon position $\langle z_n \rangle$ (*solid squares*, in Å) for all chains.

It is also important to consider the mean square travel,

$$\Delta_n^2 \equiv \langle D_n^2 \rangle = \int D^2 p_n(D) \mathrm{d}D. \tag{7}$$

These quantities are shown in Fig. 4.

To consider averages of the order parameter S_n, it is necessary to consider another variable in the probability distribution functions. Fig. 2 shows the local axis z' that goes through carbons $n-1$ and $n+1$ and makes the angle θ with the bilayer normal along z. Rotation around the local axis z' is measured by the angle ψ. A straightforward calculation (see Eq. 3a; Nagle, 1993) gives the combined order parameter S_n of both deuterons on the nth methylene corresponding to this orientation:

$$S_n = \tfrac{1}{2}\sin^2\psi - \tfrac{1}{2}(1 + \sin^2\psi)\cos^2\theta, \tag{8}$$

where $\cos\theta = D_n/D_M$ involves only the previous variable D_n and the maximum travel along the z axis, which is $D_M = 2.54$ Å for undistorted saturated hydrocarbon chains. The additional variable is $\sin^2\psi$, for which average values for narrow ranges of D are given by

$$\langle \sin^2\psi_n(D) \rangle \equiv \int \sin^2\psi_n(D) p_n(D, \psi) \mathrm{d}\psi. \tag{9}$$

Fig. 5 shows values of $\langle \sin^2\psi_n(D) \rangle$ for two values of n for the upper and lower monolayers separately. The fact that $\langle \sin^2\psi_n(D) \rangle$ is closer to 0.4 for the upper monolayer and closer to 0.5 for the lower monolayer indicates incomplete thermal equilibration. This variable was considered by van der Ploeg and Berendsen (1983), with the suggestion that $\langle \sin^2\psi_n(D) \rangle$ was less than 0.5.

After integration of Eq. 8 over ψ, the average order parameter

$$\langle S_n(D) \rangle = \int S_n(D, \psi) p_n(D, \psi) \mathrm{d}\psi \tag{10}$$

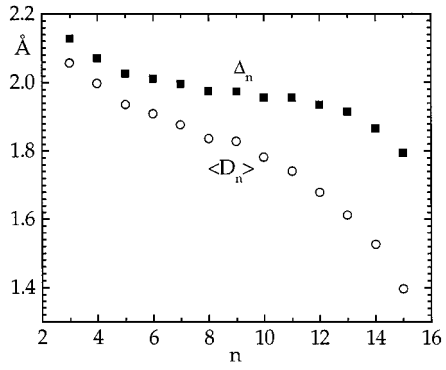

FIGURE 4 Simulated data for Δ_n (■) and $\langle D_n \rangle$ (○) averaged for all chains.

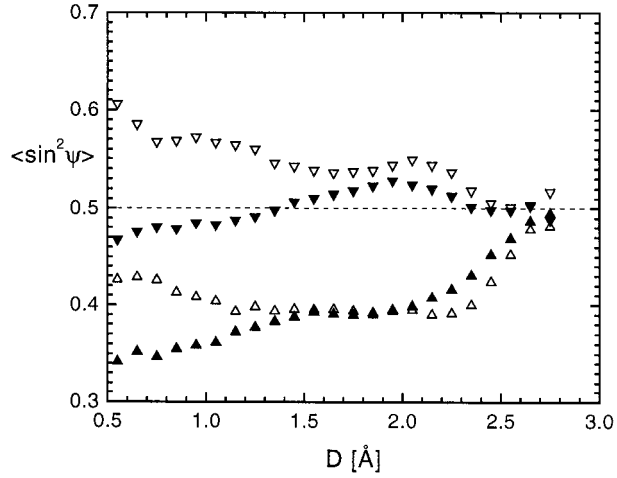

FIGURE 5 Simulated $\langle \sin^2\psi_n(D) \rangle$ as a function of D for $n = 6$ (△, ▽) and $n = 12$ (▲, ▼) averaged over sn-1 and sn-2 chains. △, ▲, The upper monolayer; ▽, ▼, the lower monolayer.

becomes

$$\langle S_n(D) \rangle = \tfrac{1}{2}\langle \sin^2\psi_n(D) \rangle - \tfrac{1}{2}(1 + \langle \sin^2\psi_n(D) \rangle)(D/D_M)^2. \tag{11}$$

Fig. 6 shows simulation results that have been binned according to the values of D, for both the upper monolayer and the lower monolayer. Fig. 6 also shows theoretical curves for $\langle S_n(D) \rangle$ that were obtained from Eq. 11 for different values of $\langle \sin^2\psi_n(D) \rangle$. Clearly, the curve using a random distribution $\langle \sin^2\psi_n(D) \rangle = 0.5$ fits quite well for 1.5 Å $< D < D_M$, where the counting statistics are good (see $p(D)$ curves in Fig. 6), and even reasonably well for smaller values of D, where the statistics are much poorer, because there are few groups with large negative values of D; such

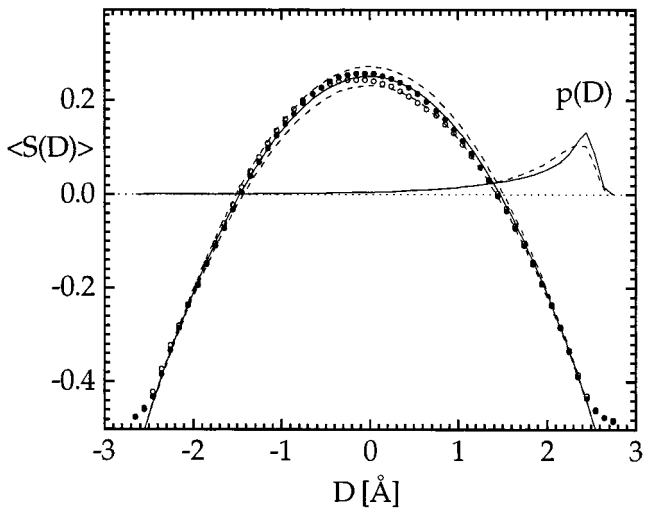

FIGURE 6 Simulated $\langle S(D) \rangle$ averaged over all carbons in the lower monolayer (●) and the upper monolayer (○) are compared to Eq. 11. ——, the result for $\langle \sin^2\psi(D) \rangle = 0.50$; – – –, the results for 0.46 and 0.54. The curves labeled $p(D)$ show the probability distribution function for the lower monolayer (——) and the upper monolayer (– – –).

groups have been called *upturns* (Nagle, 1993). (The occurrence of values of D greater than D_M is due to molecular distortions caused by thermal fluctuations and intermolecular interactions.)

Based on Fig. 6, we now adopt the approximation $\langle \sin^2\psi_n(D)\rangle = 0.5$. Then,

$$\langle S_n(D)\rangle = \frac{1}{4}\left[1 - 3\left(\frac{D}{D_M}\right)^2\right]. \quad (12)$$

Further integration over D yields the usual average order parameters,

$$\langle S_n\rangle = \frac{1}{4}\left[1 - 3\left(\frac{\Delta_n}{D_M}\right)^2\right]. \quad (13)$$

The consistency of Eq. 13 is tested in Fig. 7, which shows that the results for the direct simulation of S_n are very close to the values of $\langle S_n\rangle$ obtained from Eq. 13 and direct simulation of Δ_n.

Equation 12 emphasizes that $\langle S_n(D)\rangle$ is a quadratic function of D, so that $\langle S_n\rangle$ is a function of $\langle D_n^2\rangle = \Delta_n^2$ rather than of $\langle D_n\rangle^2$. This just reflects the well-known fundamental fact that the order parameter is a second-order Legendre polynomial, whereas the travel is first order. One might nevertheless consider approximating $\langle D_n\rangle$ by Δ_n:

$$\langle D_n\rangle \approx \Delta_n = D_M\sqrt{\frac{1 - 4\langle S_n\rangle}{3}}, \quad (14)$$

where the equality comes from inverting Eq. 13. This is a poor approximation, as shown first by Fig. 4, which compares $\langle D_n\rangle$ and Δ_n, and, more importantly, by the lower curved line in Fig. 1. The reason for this is that Δ_n is generally greater than $\langle D_n\rangle$ because $p_n(D)$ has a nonzero width σ_n:

$$\sigma_n^2 \equiv \langle (D_n - \langle D_n\rangle)^2\rangle = \Delta_n^2 - \langle D_n\rangle^2. \quad (15)$$

Because $\sigma_n^2 > 0$, the methylenic travel given by Eq. 14 is overestimated. This is a basic mathematical problem. Upturns, defined as groups with $D < 0$, are involved, in so far as they broaden the distribution and increase σ_n for larger n, as indicated by the increasing difference between Δ_n and $\langle D_n\rangle$ in Fig. 4, but they are not the fundamental problem.

To calculate D_n from Δ_n, we need a way to estimate the mean square deviation σ_n. Such a way is suggested by considering the distribution function $p_n(D)$ in Fig. 8, which shows that, over the most significant range of D, $p_n(D)$ behaves roughly exponentially:

$$p_n(D) = \mathcal{P}_n e^{D/\lambda_n}, \quad \text{for } D \le D_M. \quad (16)$$

The relevant parameter is the decay length λ_n, and the parameter $\mathcal{P}_n$ is just a normalization factor. It may be noted that the decay length λ_n increases with the carbon number n in Fig. 8. Furthermore, secondary peaks in $p_n(D)$ occur near maximally negative values of D; these are clearly deviations from the functional form in Eq. 16. These are due to upturns, which are infrequent for small n, but become more numerous, although they are still less than 10%, for large n.

Assuming Eq. 16, we have

$$\Delta_n^2 = \mathcal{P}_n\int_{-D_M}^{D_M} D^2 e^{D/\lambda_n} dD = D_M^2 + 2\lambda_n^2 - 2\lambda_n D_M \tanh\frac{D_M}{\lambda_n} \quad (17)$$

$$\langle D_n\rangle = \mathcal{P}_n\int_{-D_M}^{D_M} D e^{D/\lambda_n} dD = D_M \tanh\frac{D_M}{\lambda_n} - \lambda_n. \quad (18)$$

With no further approximation, Eq. 17 can be solved numerically to give λ_n. However, analytical expressions are more appealing, and for this reason we extend the above integrals from $-D_M$ to $-\infty$ (valid for $\lambda_n \ll D_M$). Then,

$$\Delta_n^2 = D_M^2 + 2\lambda_n^2 - 2\lambda_n D_M \quad (19)$$

$$\langle D_n\rangle = D_M - \lambda_n \quad (20)$$

$$\sigma_n^2 = \lambda_n^2. \quad (21)$$

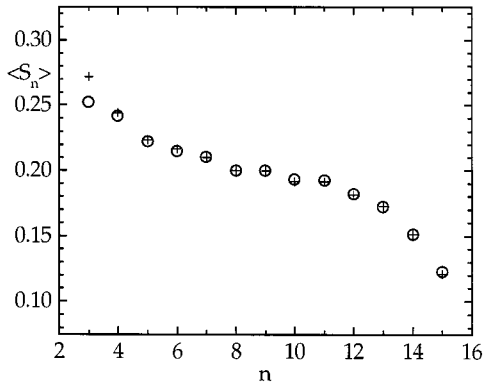

FIGURE 7 Comparison of $\langle S_n\rangle$ obtained from Eq. 13 (+) with directly simulated $\langle S_n\rangle$ (○).

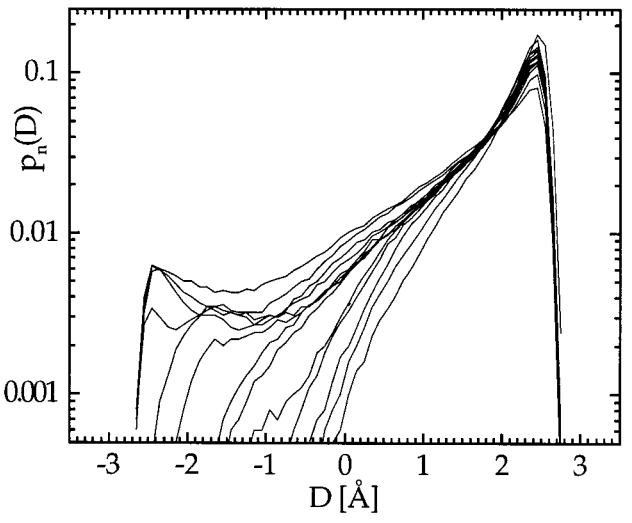

FIGURE 8 Number distributions $p_n(D)$ for $n = 3$ (smallest decay length) to $n = 15$ (largest decay length).

Solving for λ_n in terms of Δ_n yields

$$\sigma_n = \lambda_n = \tfrac{1}{2}(D_M - \sqrt{2\Delta_n^2 - D_M^2}) \quad (22)$$

$$\langle D_n \rangle = \tfrac{1}{2}(D_M + \sqrt{2\Delta_n^2 - D_M^2}). \quad (23)$$

The square roots above are imaginary for $\Delta_n^2 < D_M^2/2$, and this sets a limit to the applicability of Eq. 23. For such small values of Δ_n^2 most of our previous approximations are questionable. In this limit, λ_n would be on the order of $D_M/2$, meaning a broad $p_n(D)$, which is poorly modeled by Eq. 16. For carbons at the end of the chain we may expect more complicated probability distributions. As shown by the simulations, the maximum in $p_n(D)$ deviates from D_M, and peaks at negative values of D start to develop because of upturns. However, for another comparison note that a constant (isotropic) distribution between $-D_M$ and D_M gives $\Delta_n^2 = D_M^2/3$; this is only a little less than the applicability limit, so the restriction $\Delta_n^2 > D_M^2/2$ is not so severe.

Assuming that the approximations hold, we can combine Eq. 23 with Eq. 13 to give our final expression of $\langle D_n \rangle$ as a function of $\langle S_n \rangle$:

$$\frac{\langle D_n \rangle}{D_M} = \frac{1}{2}\left(1 + \sqrt{\frac{-8\langle S_n \rangle - 1}{3}}\right), \quad (24)$$

where the corresponding limiting point of applicability is $\langle S_n \rangle = -1/8 = -0.125$. In these simulations, only $\langle S_{15} \rangle$ is close to this limit, as can be seen in Fig. 1, which plots Eq. 24 and compares to direct simulations. The usual way to compare is shown in Fig. 9, which compares the direct simulation results for $\langle D_n \rangle$ with those obtained from Eq. 24, using the simulation results for $\langle S_n \rangle$. These two results deviate for large n because the exponential form used in Eq. 16 is not accurate for methylenes near the terminal methyl, as shown in Fig. 8. Moreover, the approximation $\lambda_n \ll D_M$ used to obtain Eq. 19 breaks down for $n > 10$. The standard Eq. 1, shown in Fig. 9, deviates at both high n and low n.

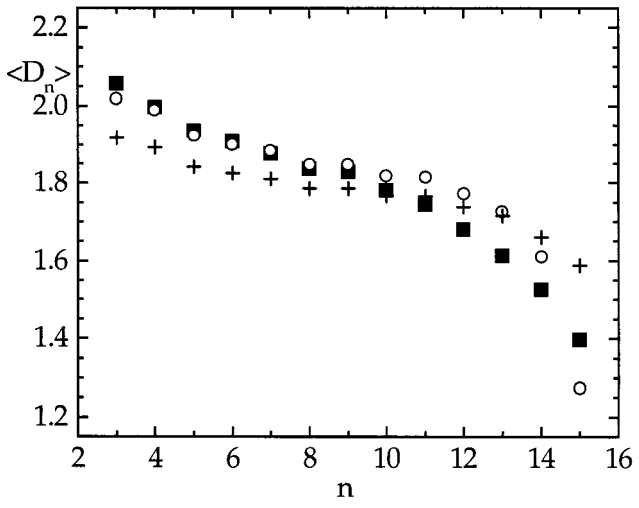

FIGURE 9 Solid squares show directly simulated $\langle D_n \rangle$ (in Å). Estimates using Eq. 24 are shown with open circles, and the + symbols show estimates from Eq. 1.

AVERAGE CHAIN LENGTH

In this section we extend the results for methylenic travel to longer segments of the hydrocarbon chain. Of particular interest is the average length of the entire hydrocarbon chain. This length is a little awkward conceptually because it should include a poorly defined piece beyond the terminal methyl $n = 16$. A more precisely defined length will be called L_C^*, which is defined to be the average distance along the bilayer normal between the first methylene carbon and the terminal methyl carbon:

$$\langle L_C^* \rangle \equiv \langle z_2 \rangle - \langle z_{16} \rangle. \quad (25)$$

For the present simulation $\langle z_2 \rangle = 13.78$ Å and $\langle z_{16} \rangle = 1.32$ Å, giving $\langle L_C^* \rangle = 12.46$ Å. To estimate the average full length of the chain $\langle L_C \rangle$, we first add 0.547 Å to $\langle L_C^* \rangle$ to account for half of the projected distance from the first methylene to the carbonyl carbon. We then add $1.5\langle z_{15} - z_{16} \rangle$ to account for the extra length of the terminal methyl; the rationale for this is that the terminal methyl volume V_{CH_3} is about twice as large as methylene volume V_{CH_2} (Nagle and Wiener, 1988; Petrache et al., 1997). The estimated chain length obtained directly from the simulations is then $\langle L_C \rangle = 13.99$ Å.

To obtain chain lengths from NMR requires using either our new Eq. 24 or the conventional Eq. 1. We first report results for $\langle L_C^* \rangle$, which are simply obtained by summing $\langle D_n \rangle$ over all odd values of n from 3 to 15. The result from Eq. 24 is $\langle L_C^* \rangle = 12.49$ Å, and the result from Eq. 1 is $\langle L_C^* \rangle = 12.42$ Å. Despite the deviations in Fig. 9, both approximate results from $\langle S_n \rangle$ agree well with the direct result, $\langle L_C^* \rangle = 12.46$ Å. However, the result from our new Eq. 24 uses $\langle D_{15} \rangle$, which is right at the limit where the square root becomes imaginary, and so it probably gives a too small value, which happens to compensate for the positive deviations in Fig. 9. The conventional result involves even more accidental cancellations, as shown in Fig. 9.

We next estimate the full chain length $\langle L_C \rangle$, using

$$\langle L_C \rangle = (1/2)\sum_{n=2}^{15}\langle D_n \rangle + \langle D_{15} \rangle, \quad (26)$$

where $\langle D_2 \rangle$ was estimated by $\langle D_3 \rangle$, and where the extra $\langle D_{15} \rangle$ term estimates the extra contribution from the $n = 16$ terminal methyl end. These conventions are identical to those used two paragraphs above to obtain $\langle L_C \rangle = 13.99$ Å. The value obtained using Eq. 26 and the direct simulation results for $\langle D_n \rangle$ is $\langle L_C \rangle = 14.01$ Å. Using Eq. 24 and the simulated $\langle S_n \rangle$ yields $\langle L_C \rangle = 14.0$ Å, and using Eq. 1 yields $\langle L_C \rangle = 14.1$ Å. It therefore appears that, for these simulations, either the old or the new formula gives good estimates of $\langle L_C \rangle$.

It is also worth considering such quantities for different simulations. Berger et al. (1997) reported relative positions of C_4, C_9, and C_{14} for three different simulation runs of fluid-phase DPPC. They also reported the order parameters for each run. Our test of Eq. 24 for their best run (no. 2) is presented in Table 1. Feller et al. (1997) reported similar results, which are also used to test Eq. 24 in Table 1. Comparison values are also shown for the present simulation. In the plateau region, $n = 4$–9, the new Eq. 24 gives superior values for the travel when compared to the conventional Eq. 1. In the extended region, $n = 4$–14, the conventional Eq. 1 does better, but the relative error in Eq. 24 is less than 3%.

HYDROCARBON THICKNESS

To obtain a better understanding of chain packing, it is especially interesting to compare the average length of the hydrocarbon chains $\langle L_C \rangle$ with half the thickness of the hydrocarbon chain region $\langle D_C \rangle$. The latter quantity is defined by

$$\langle D_C \rangle \equiv \frac{V_C}{\langle A \rangle}, \tag{27}$$

where $\langle A \rangle$ is the average area per molecule, which is 61.8 Å^2 in this simulation, and V_C is the volume of both hydrocarbon chains, which has been determined to be 862 Å^2 for this simulation, using the procedure of Petrache et al. (1997). Using Eq. 27 then gives $\langle D_C \rangle = 13.95$ Å. The consistency of this procedure is indicated in Fig. 10, which shows where $\langle D_C \rangle$ falls on a profile of the hydrocarbon probability distribution function:

$$P_{HC}(z) = \sum_{n=2}^{16} V_n p_n(z), \tag{28}$$

where $V_n = V_{CH_2}$ for $n = 2$–15 and $V_{16} = V_{CH_3}$. (The flatness of the hydrocarbon probability distribution in the region -8 Å $< z < 8$ Å, where the headgroups and water are absent, is the criterion used to obtain the ratio V_{CH_3}/V_{CH_2}, and then V_{CH_2} comes from the simulated density of hydrocarbon in this region (Petrache et al., 1997).) Fig. 10 shows

TABLE 1 Comparison of partial chain lengths from different simulations (direct) and estimates using Eqs. 24 and 1

	Tu et al.	Feller et al.	Berger et al.
$\langle z_4 \rangle - \langle z_9 \rangle$ (Å)			
Direct	4.76	4.80	4.65
Eq. 24	4.73	4.70	4.63
Eq. 1	4.54	4.52	4.41
$\langle z_4 \rangle - \langle z_{14} \rangle$ (Å)			
Direct	9.00	8.80	8.50
Eq. 24	9.19	8.85	8.75
Eq. 1	8.92	8.73	8.55

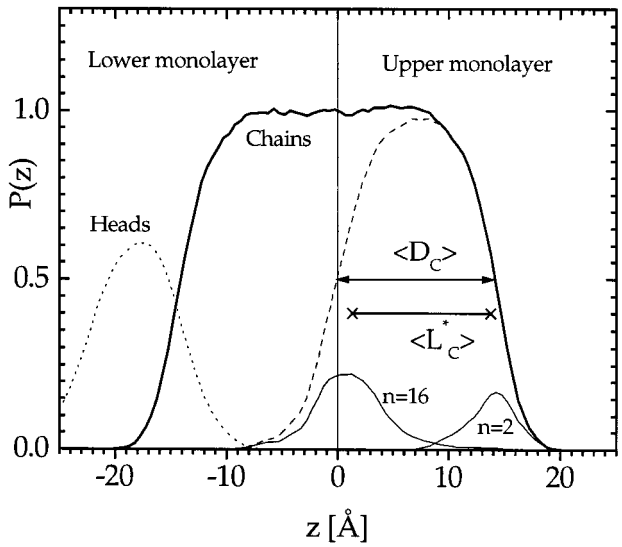

FIGURE 10 Probability distribution of hydrocarbon chains, using $V_{CH_2} = 26.94$ Å^3 and $V_{CH_3} = 53.80$ Å^3. *Thick solid line:* Contribution from both monolayers. *Dashed line:* Upper monolayer only. *Crosses:* Average positions $\langle z_2 \rangle$ and $\langle z_{16} \rangle$. *Thin solid lines:* Probability distributions for $n = 2$ and $n = 16$ (terminal methyl) in the upper monolayer. *Dotted line:* Headgroup (including carbonyls, glycerol, and phosphatidylcholine) probability distribution in the lower monolayer.

where $\langle z_{16} \rangle$ and $\langle z_2 \rangle$ fall for the upper monolayer, as well as the probability distributions for these carbons.

A major and unexpected result of this simulation is that $\langle L_C \rangle$ determined in the previous section is numerically very close to $\langle D_C \rangle$. Although this result has often been implicitly assumed, it is not a priori correct, as emphasized by Nagle (1993). This point is illustrated in Fig. 11, which shows three caricatures, each of which packs four chains, two in each monolayer. Because of this simplicity, the average chain length in these caricatures is just $\langle L_C \rangle = (L_1 + L_2)/2$. The distinguishing feature of model I is early chain termination of chain 1, so $L_1 < D_C$, which makes $\langle L_C \rangle$ less than $\langle D_C \rangle$. Model I has decreasing order with increasing n because the terminal methyl end of chain 2 is more disordered than the carbonyl end, whereas chain 1 is equally disordered along the chain. The distinguishing feature of model II is interdigitation across the midplane, which makes $\langle L_C \rangle$ greater than $\langle D_C \rangle$, and which results in increasing order with increasing n. Because the experimental order parameter decreases with increasing n, model I is clearly superior to model II. However, for this simulation, which has $\langle L_C \rangle = \langle D_C \rangle$, model I must be wrong. This has led us to propose model III in Fig. 11. Model III has both early chain terminations and interdigitation, the order parameter decreases with increasing n, and $\langle L_C \rangle = \langle D_C \rangle$, so model III is the best model for characterizing these simulations.

AREA PER MOLECULE

For the present simulation, $\langle A \rangle = 61.8$ Å^2 is obtained simply by dividing the total area of either monolayer by the number

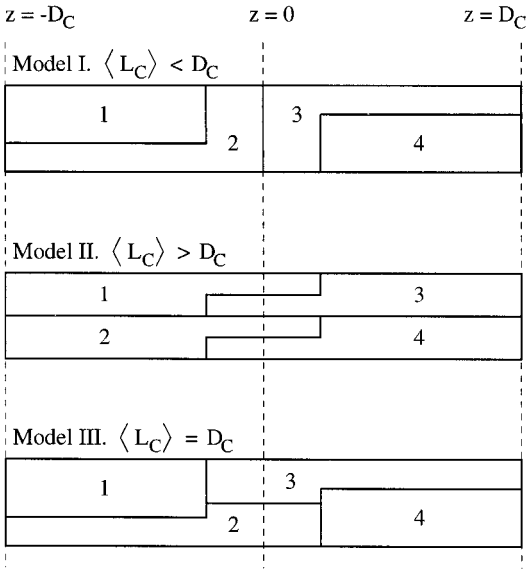

FIGURE 11 Three simplified caricatures, labeled I, II, and III, of packing geometry. For simplicity, only four chains are shown for each model, but more realistic distribution functions can easily be envisioned. The four numbered boxes represent the volumes occupied by each of the four chains. Ignoring upturns, the chains start from the headgroup end at $z = \pm D_C$ and end in the $z = 0$ region. The order parameter is larger when the vertical cross section of the box is smaller. Note that these caricatures are not meant to designate sn-1 versus sn-2 chains.

of lipid molecules in that layer. The issue is whether this value of $\langle A \rangle$ can be obtained using the NMR order parameters. For this simulation, which conforms to chain packing model III in Fig. 11, this is easily accomplished using Eq. 27 and the preceding estimates of $\langle L_C \rangle$ for $\langle D_C \rangle$; this gives the satisfactory result $\langle A \rangle = 862$ Å^3/14.0 Å $= 61.6$ Å^2.

Because model III may not apply to all bilayers, it is also of interest to evaluate a different way to obtain $\langle A \rangle$ that uses only the low carbon numbers in the so-called plateau region. The usual formula for this is

$$\langle A_n \rangle = 4V_{CH_2}/\langle D_n \rangle, \tag{29}$$

where $V_{CH_2} = 26.94$ Å^3 is the volume per methylene group in this simulation, and the factor of 4 accounts for two chains per lipid and for a factor of 2 in D_n, which, as defined, is really twice the travel per methylene. Fig. 12 shows the results from Eq. 29 for $\langle A_n \rangle$ versus n. These values are all too low in the plateau region $n = 3$–8.

There is a basic mathematical flaw in Eq. 29, which is corrected by taking the average of the instantaneous area, $A_n = 4V_{CH_2}/D_n$; this gives (Brown, 1996)

$$\langle A_n \rangle = 4V_{CH_2}\langle 1/D_n \rangle. \tag{30}$$

The relation

$$\left\langle \frac{1}{D_n} \right\rangle = \left\langle \frac{1}{\langle D_n \rangle + (D_n - \langle D_n \rangle)} \right\rangle \approx \frac{1}{\langle D_n \rangle} + \frac{\langle (D_n - \langle D_n \rangle)^2 \rangle}{\langle D_n \rangle^3}$$

$$= \frac{1}{\langle D_n \rangle} \frac{\langle D_n^2 \rangle}{\langle D_n \rangle^2} \tag{31}$$

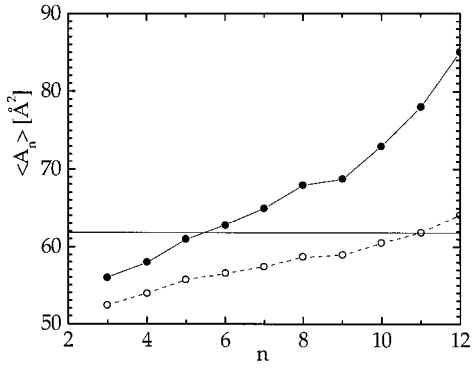

FIGURE 12 Estimate of A_n, using Eq. 29 (○) and Eq. 31 (●). The horizontal line indicates the actual simulated area $A = 61.8$ Å.

shows that $\langle 1/D_n \rangle$ is larger than $1/\langle D_n \rangle$ because $\langle D_n^2 \rangle$ is larger than $\langle D_n \rangle^2$. The order parameters are used to obtain $\langle D_n \rangle$ (Eq. 24) and $\langle D_n^2 \rangle$ (Eq. 13). Fig. 12 shows that $\langle A_n \rangle$ obtained using Eq. 31 in the plateau region is in reasonable agreement with the actual $\langle A \rangle$ if one averages $\langle A_n \rangle$ over a plateau region defined to be $n = 3$–8.

Although the average over $n = 3$–8 gives the correct $\langle A \rangle$, it is important to understand why there is a slope in $\langle A_n \rangle$. Fig. 3 shows the average positions $\langle z_n \rangle$, and Fig. 10 shows that there is a significant fraction of the headgroups at values of $\langle z_n \rangle$ for $n = 3$–5. Because these headgroups take up volume and area that is not accounted for by Eq. 30, this accounts for the smaller values of $\langle A_n \rangle$ in Fig. 12. The probability of headgroups decreases sharply at $\langle z_n \rangle$ for larger values of n, but Fig. 10 shows that the probability of terminal methyls increases. These early chain terminations mean that there is more total area that is shared by fewer chains, which is also not taken into consideration in Eq. 30; this accounts for the larger values of $\langle A_n \rangle$ in Fig. 12. The chain termination consideration was known previously (Nagle, 1993), and this motivated considering only the plateau region. This simulation shows that the headgroups must also be considered and that there is hardly any region that can be said to be free of both artifacts. Although one can obviously define a plateau region that works for this simulation, it is not clear if this will be universal for other simulations or for real lipid bilayers.

DISCUSSION

The main result in this paper has been the development of Eq. 24, which gives a new approximation for obtaining methylenic travel $\langle D_n \rangle$ from NMR order parameters $\langle S_n \rangle$. As shown in Fig. 1, this formula fits the results of this simulation better than older approximations. Of course, the derivation of Eq. 24 used detailed results of the probability distribution shown in Fig. 8, and so Eq. 24 can only be expected to be as good as the simulation. There are results that show that the ψ angle for the chains did not come to equilibrium in the simulation (Fig. 5), although this seems to make only a minor difference in the tests in Figs. 6 and

7. Furthermore, the results from the different simulation of Berger et al. (1997) are in quantitatively good agreement, and results for chain fragments from three simulations (Table 1) are in good agreement. Finally, the use of Eq. 24 to predict the chain length is in excellent agreement with the actual chain length $\langle L_C \rangle$. Nevertheless, Eq. 24 is a sufficiently radical departure from standard practice that it should be tested on new and independent simulations of bilayers.

A major surprise is the simulation result that half the mean thickness of the hydrocarbon chain region $\langle D_C \rangle$ is numerically close to the average chain length $\langle L_C \rangle$. Even though this has been an implicit assumption in NMR studies (Schindler and Seelig, 1975; Thurmond et al., 1991), it is not a priori necessary. In particular, the occurrence of early chain terminations (model I in Fig. 11) would, by itself, require $\langle L_C \rangle$ to be smaller than $\langle D_C \rangle$. However, the effect of early chain terminations appears to be canceled by the effect of interdigitation, as illustrated by model III in Fig. 11.

The result $\langle D_C \rangle = \langle L_C \rangle$ leads to a simple way of estimating area per molecule $\langle A \rangle$ using Eq. 27. This reproduces the actual $\langle A \rangle$ in this simulation. This would be a very nice result, except that it leads to a disagreement. This procedure gave $\langle A_{DPPC} \rangle = 71.7$ Å^2 for DPPC at 50°C (Thurmond et al., 1991). Of course, the NMR study used the older Eq. 1, but we have shown in the third section of this paper that, for this simulation, the significant deviations in this formula for individual carbons n (see Fig. 9) average out, so that the total chain length $\langle L_C \rangle$ is quite well approximated by either Eq. 1 or Eq. 24. Therefore, the results in this paper support the method used by Thurmond et al. (1991). However, their result for $\langle A_{DPPC} \rangle$ is 14% larger than the value $\langle A_{DPPC} \rangle = 62.9$ Å^2 obtained from x-ray studies (Nagle et al., 1996).

We have examined this disagreement further for DMPC. Two recent studies have agreed that $\langle A_{DMPC} \rangle = 59.6 \pm 0.5$ Å^2 for DMPC at $T = 30$°C. Our study (Petrache et al., 1998) used the same x-ray method as used for DPPC. The other study (Koenig et al., 1997) combined x-ray and NMR results, but only changes in $\langle L_C \rangle$ were obtained from NMR order parameters because of the concern that no formula was adequate to give absolute numbers. Because these two studies used quite different methods and assumptions, we believe that the common value of $\langle A_{DMPC} \rangle$ that was obtained is a reliable benchmark for testing present NMR formulae.

We used the order parameters obtained by Koenig et al. (1997) at $T = 30$°C in Eqs. 24 to estimate $\langle D_n \rangle$, which were then added up, using Eq. 26, to obtain $\langle L_C \rangle = 11.95$ Å. We then assumed that $\langle D_C \rangle = \langle L_C \rangle$ in Eq. 27. We also used $V_C = 782$ Å^3, which was obtained by subtracting the headgroup volume $V_H = 319$ Å^3 (Sun et al., 1994) from the lipid volume $V_L = 1101$ Å^3 (Nagle and Wilkinson, 1978). The result obtained with Eq. 27 is $\langle A_{DMPC} \rangle = 65.4$ Å^2, 10% higher than obtained previously. Using the conventional Eq. 1 instead of Eq. 24 reduces this marginally to $\langle A_{DMPC} \rangle = 64.3$ Å^2. Using the plateau method described in the previous section and averaging $\langle A_n \rangle$ over $n = 3$–8 gives $\langle A_{DMPC} \rangle = 65.2$ Å^2. Therefore, for DMPC, the methods that work so well for this simulation give values for $\langle A_{DMPC} \rangle$ that are too large.

The resolution of this disagreement may involve several issues. The first and most obvious is whether this simulation is misleading. Clearly, other simulations should be analyzed with NMR interpretation in mind. This involves testing Eq. 24, which also has implications for changes in chain length used by Koenig et al. (1997). It also involves testing $\langle L_C \rangle = \langle D_C \rangle$ and Eq. 27. Of course, it is possible that all simulations will give the same but not the correct answer to the issue of equality of $\langle L_C \rangle$ and $\langle D_C \rangle$ because chain packing, as contrasted with chain conformational order, does not equilibrate in the nanosecond time range. In this regard NMR studies indicate slower, collective motions (Nevzorov et al., 1998). It may be noted that the simulation studied in this paper was a particularly long one, although a new hybrid Monte Carlo/molecular dynamics method may, in the future, overcome some equilibration barriers (Clark et al., 1999). Perhaps one might also question whether, because of subtle effects of relative time scales (Brown, 1996), NMR may not measure the same instantaneous order parameters that are obtained straightforwardly from simulations. Measured NMR order parameters that are only ~10% too low would also account for the disagreement.

We have not yet achieved the desired goal of being able to advocate an analysis method that can use NMR order parameters to quantitate bilayer structure. Nevertheless, flaws in previous methods have been elucidated, and the issues have been drawn more precisely. Further simulations and further analysis of already completed and ongoing simulations should be performed to help resolve the remaining issues.

We thank Klaus Gawrisch for supplying the order parameters for DMPC, Doug Tobias for discussions, and Michael Klein for support of KT (NIH grant GM 40712) when her initial work was performed. NIH grant GM44976 is gratefully acknowledged for support of HIP and JFN.

REFERENCES

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

Brown, M. F. 1996. Membrane structure and dynamics studied with NMR spectroscopy. *In* Biological Membranes. K. Merz and B. Roux, editors. Birkhäuser, Boston.

Clark, M. M., S.-W. Chiu, E. Jakobsson, H. L. Scott, and S. Subramaniam. 1999. Computer simulations of DPPC and DPPC-cholesterol bilayers. *Biophys. J.* 74:A12.

De Young, L. R., and K. A. Dill. 1988. Solute partitioning into lipid bilayer membranes. *Biochemistry.* 27:5281–5289.

Feller, S. E., R. M. Venable, and R. W. Pastor. 1997. Computer simulation of a DPPC phospholipid bilayer: structural changes as a function of molecular surface area. *Langmuir.* 13:6555–6561.

Koenig, W., H. H. Strey, and K. Gawrisch. 1997. Membrane lateral compressibility determined by NMR and x-ray diffraction: effect of acyl chain polyunsaturation. *Biophys. J.* 73:1954–1966.

Nagle, J. F. 1993. Area/lipid of bilayers from NMR. *Biophys. J.* 64: 1476–1481.

Nagle, J. F., and M. C. Wiener. 1988. Structure of fully hydrated bilayer dispersions. *Biochim. Biophys. Acta.* 942:1–10.

Nagle, J. F., and D. A. Wilkinson. 1978. Lecithin bilayers: density measurements and molecular interactions. *Biophys. J.* 23:159–175.

Nagle, J. F., R. Zhang, S. Tristram-Nagle, W.-J. Sun, H. I. Petrache, and R. M. Suter. 1996. X-ray structure determination of fully hydrated L_α phase dipalmitoylphosphatidylcholine bilayers. *Biophys. J.* 70: 1419–1431.

Nevzorov, A. A., T. P. Trouard, and M. F. Brown. 1998. Lipid bilayer dynamics from simultaneous analysis or orientation and frequency dependence of deuterium spin-lattice and quadrupolar order relaxation. *Phys. Rev. E.* 58:2259–2281.

Petrache, H. I., S. E. Feller, and J. F. Nagle. 1997. Determination of component volumes from simulations. *Biophys. J.* 70:2237–2422.

Petrache, H. I., S. Tristram-Nagle, and J. F. Nagle. 1998. Fluid phase structure of EPC and DMPC bilayers. *Chem. Phys. Lipids*. 95:83–94.

Schindler, H., and J. Seelig. 1975. Deuterium order parameters in relation to thermodynamic properties of a phospholipid bilayer. *Biochemistry*. 14:2283–2287.

Sun, W.-J., R. M. Suter, M. A. Knewtson, C. R. Worthington, S. Tristram-Nagle, R. Zhang, and J. F. Nagle. 1994. Order and disorder in fully hydrated unoriented bilayers of gel phase DPPC. *Phys. Rev. E.* 49: 4665–4676.

Thurmond, R. L., S. W. Dodd, and M. F. Brown. 1991. Molecular areas of phospholipids as determined by ^{2}H NMR spectroscopy. *Biophys. J.* 59:108–113.

Tieleman, D. P., S. J. Marrink, and H. J. C. Berendsen. 1997. A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochim. Biophys. Acta.* 1331:235–270.

Tobias, D. J., K. Tu, and M. L. Klein. 1997. Atomic-scale molecular dynamics simulations of lipid membranes. *Curr. Opin. Colloid Interface Sci.* 2:15–26.

Tu, K., D. J. Tobias, and M. L. Klein. 1995. Constant pressure and temperature molecular dynamics simulation of a fully hydrated liquid crystal phase dipalmitoylphosphatidylcholine bilayer. *Biophys. J.* 69: 2558–2562.

van der Ploeg, P., and H. J. C. Berendsen. 1983. Molecular dynamics of a bilayer membrane. *Mol. Phys.* 49:233–248.