University of Groningen

Atomistic and Coarse Grain Topologies for the Cofactors Associated with the Photosystem II Core Complex

de Jong, Djurre H.; Liguori, Nicoletta; van den Berg, Tom; Arnarez, Clement; Periole, Xavier; Marrink, Siewert J.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

# Atomistic and Coarse Grain Topologies for the Cofactors Associated with the Photosystem II Core Complex

Djurre H. de Jong,[†,§] Nicoletta Liguori,[‡] Tom van den Berg,[†,‖] Clement Arnarez,[†] Xavier Periole,*[,†] and Siewert J. Marrink*[,†]

[†]Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

[‡]Department of Physics and Astronomy and Institute for Lasers, Life and Biophotonics, Faculty of Sciences, VU University Amsterdam, De Boelelaan 1081, 1081 HV, Amsterdam, The Netherlands

Ⓢ Supporting Information

ABSTRACT: Electron transfers within and between protein complexes are core processes of the electron transport chains occurring in thylakoid (chloroplast), mitochondrial, and bacterial membranes. These electron transfers involve a number of cofactors. Here we describe the derivation of molecular mechanics parameters for the cofactors associated with the function of the photosystem II core complex: plastoquinone, plastoquinol, heme b, chlorophyll A, pheophytin, and β-carotene. Parameters were also obtained for ubiquinol and ubiquinone, related cofactors involved in the respiratory chain. Parameters were derived at both atomistic and coarse grain (CG) resolutions, compatible with the building blocks of the GROMOS united-atom and Martini CG force fields, respectively. Structural and thermodynamic properties of the cofactors were compared to experimental values when available. The topologies were further tested in molecular dynamics simulations of the cofactors in their physiological environment, e.g., either in a lipid membrane environment or in complex with the heme binding protein bacterioferritin.
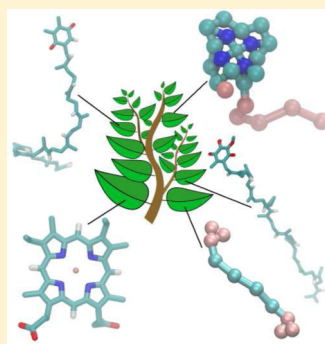
■ INTRODUCTION

Most life on earth is dependent on sunlight for its energy. The conversion of sunlight into chemical energy happens in the electron transport chain located in the chloroplasts of green plants, algae, and cyano-bacteria. In these organisms, the conversion is taking place in the thylakoid membrane: a densely packed membrane mostly consisting of thylakoid lipids and four large protein complexes.[1−3] The complexes, photosystems I and II (PSI and PSII), cytochrome b_6f, and F-ATPase perform the light conversion. PSII absorbs photons to oxidize water and reduce the cofactor plastoquinone into plastoquinol. Plastoquinol diffuses to cytochrome b_6f where it is reoxidized, releasing energy to reduce plastocyanin. The reduced plastocyanin is reoxidized by PSI. The released energy, together with the energy from the uptake of another photon by PSI, is used to reduce ferredoxin, which is released to the cell stroma. During these steps, a proton gradient is created over the membrane, which is ultimately used by F-ATPase to create ATP from ADP.[4]

In the mitochondrial electron transport chain or respiratory chain, ubiquinol carries electrons from NADH oxidoreductase (complex I) to cytochrome bc1 (complex III).[5,6] The molecular structures of ubiquinone/ubiquinol are very similar to those of plastoquinone/plastoquinol. Ubiquinol is oxidized into ubiquinone by complex III, which directs the two released electrons to cytochrome c (complex IV); the two free protons are pumped out of the membrane to form a proton gradient

ultimately used by ATP synthase (complex V). Ubiquinone is then reduced back to ubiquinol by complex I.

Both the thylakoid membrane and the respiratory chain function via a complex interplay between many proteins, specific lipids, and a large number of cofactors. Apart from the electron carriers plastoquinone/ol and ubiquinone/ol, a number of other cofactors play an important role in the energy conversion taking place in these systems. These include chlorophyll A/B, pheophytin, heme, and several carotenoids such as β-carotene, zeaxanthin, lutein, and violaxanthin.

In recent years, several high resolution structures of these protein complexes together with their cofactors have been published, e.g., PSII,[7,8] PSI,[9] and respiratory chain complexes I−IV.[10−13] The availability of high resolution structures has opened the way for molecular dynamics (MD) simulation studies to investigate the dynamics of the complexes, both at full atomistic resolution[14−21] as well as using CG models.[22,23]

Atomistic parameters for some cofactors compatible with the OPLS or AMBER force fields are available.[24−29] Here we derive two sets of parameters for the cofactors associated with PSII core complex function: heme b (HEM), chlorophyll A (CLA), pheophytin (PHO), plastoquinone (PQ9one), plastoquinol (PQ9ol), and β-carotene (BCR) (Figure 1). One set of parameters is compatible with the united-atom (UA)

Figure 1. Molecular structure of the parametrized cofactors. (A) Atomistic structures where rounded shapes encircle atoms that are mapped together to one CG bead. (B) CG structures. The colors are indicative of the bead type: red, charged; yellow, polar; green, intermediate; blue, apolar. The following molecules are shown: chlorophyll A (CLA); the CG eight bead ring topology is shown. Pheophytin (PHO) has the same topology, lacking the central magnesium. Heme b (HEM); the CG eight bead ring topology is shown. β-Carotene (BCR). Plastoquinone (PQ9one), shown with nine isoprenyl units. Plastoquinol (PQ9ol) shares the same topology, except the ester groups are reduced to hydroxyl groups. Ubiquinone (UQ10one), shown with 10 isoprenyl units. Ubiquinol (UQ10ol) shares the same topology, except the ester groups have been reduced to hydroxyl groups.

GROMOS force field.[30] The other set is compatible with the CG Martini force field.[31] The use of the GROMOS parameter set as the basis for the determination of the Martini parameter set makes them consistent with one another and appropriate for eventual multiscale studies.[32,33] In addition, we parametrized ubiquinone (UQ10one) and ubiquinol (UQ10ol), cofactors associated with the respiratory chain complexes.

The remainder of this paper is organized as follows. In the Methods section, the protocols used to derive and test both the UA GROMOS and CG Martini models are described. In the Results and Discussion section, the final topologies are described in detail and critical steps in the parametrization process are discussed. In addition, the behavior of the cofactors in their physiological environment is described. A short concluding section ends this paper.

■ METHODS

Simulation Parameters. All simulations were performed using the Gromacs simulation package,[34] version 4.5.x and 4.6.x. For the UA simulations, a time step of 2 fs was used. Following the typical GROMOS setup, a twin-range cutoff scheme was used with the neighborlist extended to 1.4 nm and updated every five steps. Electrostatic interactions were treated using the reaction-field algorithm[35] with ε_{rf} = 2, 2, 10.3, or 54 when decane, cyclohexane, octanol, or water was used as a solvent, respectively. Data reported in the main manuscript are for hydrated octanol. Dry octanol was used for comparison and is reported as Supporting Information. van der Waals

interactions were cut off after 1.4 nm. Temperature and pressure were kept constant at 303 K and 1.0 bar by coupling to an external bath[36] with coupling parameters τ_t and τ_p being 0.1 and 0.5 ps^{-1}, respectively. All bonds were constrained using the LINCS algorithm.[37] UA trajectories were mapped to CG trajectories using the reverse transformation protocol.[38]

In the CG simulations, a time step of 20 fs was used. The van der Waals interactions were smoothly shifted to zero between 0.9 and 1.2 nm, adhering to the standard Martini protocol. The electrostatic interactions were explicitly screened by ε_r = 15 and were smoothly shifted to zero between 0.0 and 1.2 nm. The neighbor list of 1.4 nm was updated every 10 steps. Pressure (1 bar) and temperature (303 K) coupling parameters, τ_t and τ_p, were set to 1.0 and 3.0 ps^{-1}. Bonds were not constrained, unless explicitly set in the topology.

Partitioning coefficients (log P) between water and octanol or water and cyclohexane were calculated using the free energy perturbation (FEP) approach as implemented in the free energy code of Gromacs. The interactions between the solute molecule and the solvent were switched off in 21 or 26 steps ("windows"), following a coupling parameter λ. In water and cyclohexane, the calculations converged with 2 ns simulations at each λ value for both CG and UA representations. In atomistic dry and hydrated octanol, 5 ns per window was needed to improve sampling. Convergence was tested for PQol, CLA, and BCR (see the Supporting Information), prompting the use of longer simulations (32 and 8 ns per window, respectively) for CLA and BCR. Convergence of the errors of such calculations is shown in the Supporting Information. The

same simulation parameters were used as described above for equilibrium simulations, except a stochastic dynamics integrator was used with a (inverse) friction coefficient of 1.0 ps. The derivative with respect to λ for every window was saved for the current and both neighboring λ values, allowing for analysis using Bennet's acceptance ratio (BAR).[39] Softcore potentials were applied to avoid high energies due to overlapping particles, using softcore parameters $\alpha = 1.3$, $\sigma = 0.47$ nm, power = 1 and $\alpha = 1.3$, $\sigma = 0.3$ nm, power = 1 for the CG and UA simulations, respectively.

The free energy of transferring a solute from water (W) to octanol/cyclohexane (Sol) was calculated as

$$\Delta G_{\text{W/Sol}} = \Delta G_{\text{W/vac}} - \Delta G_{\text{Sol/vac}} \qquad (1)$$

where $\Delta G_{\text{Sol/vac}}$ is the free energy obtained from the FEP for both solvents (Sol), which can be either octanol or cyclohexane. From this free energy, the partition coefficient (log P) can be calculated using

$$\log P = \frac{-\Delta G}{\ln(10) \cdot R \cdot T} \qquad (2)$$

where $\ln(10)$ is the natural logarithm of 10, R the gas constant, and T the simulation temperature.

Parameterization of UA Models. An initial guess for the topology of each of the cofactors was obtained using either the PRODRG-server,[40] the Automated Topology Builder,[41] or previous topologies.[30] The atom types (defining the non-bonded interactions) and bonded interactions in these topologies were adjusted on the basis of the GROMOS force field building blocks[30] and the GROMOS 53A6 definitions.[42] To obtain partial charges, first the optimized geometry of an all-atom structure (obtained from the HIC-Up server,[43] hydrogens added using the PRODRG-server[40]) was obtained using the restricted Hartree–Fock (HF) method implemented in GAMESS-US[44] with a 6-31G* basis set. The effect of using a higher level of theory on the optimized structures and the atomic partial charges was tested on PQone and PQol. The results are described in the Supporting Information. Second, using the optimized structure, atomic partial charges were calculated using the same level of theory and fitting the electric potential to a Connolly surface with a density of 71.4 points per Å^2 while constraining the total monopole and dipole of the molecules obtained in the HF calculations. For the elongated shape of CLA, this protocol resulted in unrealistic high partial charges in the (aliphatic) tail. In order to avoid this issue, we calculated the partial charges for CLA using an in-house implementation of the dipole preserving analysis[45,46] (DPA) in GAMESS-UK.[47]

The final sets of partial charges were obtained as follows: (1) since the GROMOS force field has no explicit nonpolar hydrogens, charges for CH_x groups (where x can be 1, 2, or 3) that are not part of an aromatic system or a long conjugated chain were summed and assigned to the carbon atom; (2) charges were adjusted to respect the symmetry of a molecule; (3) all charges were rounded. As an additional constraint for the partial charges of PQone, PQol, UQone, and UQol, the total charge of a prenyl unit should be zero, to allow the charges of subsequent units to be equal.

Parameterization of CG Models. The mapping of atoms to CG interaction sites and the choice of bead types was based on the building blocks of Martini v2.0, as described by Marrink et al.[48] and following the recipe described in Marrink et al.[49,50]

Although no systematic rules can be given for the choices made, we followed a few general guidelines: (1) Atoms were grouped in order to unite specific chemical groups in one bead. (2) The number of atoms in a bead should be close to four for normal type beads, and not more than three for S-type beads (which are used in Martini to represent planar structures). An exception was made for the metal centers of CLA and HEM: their larger size and mass are better modeled when represented by a single S-type bead. (3) Ring motifs should consist of at least three beads. (4) Repeated motifs should be modeled using identical topologies, and topologies should reflect the symmetry of the molecules. (5) Similar molecules should share similar topologies, and differences between molecules (e.g., the Mg in CLA that is not present in PHO) are placed at the end of the topology for simplicity. In this way, the numbering of beads and bonded interactions are the same for similar molecules.

In general, CG interaction sites contain no (partial) charges. Exceptions are the central bead in HEM and CLA (+0.4 e and +1.0 e for Fe and Mg, respectively) and the four surrounding beads (−0.1 e and −0.25 e for HEM and CLA, respectively). This strategy was adopted in order to mimic the strong polar nature of the metal center. Note however that the polarizing effect of these (partial) charges is limited in Martini, since other polar groups do not have partial charges. It will affect interactions with charged beads such as ions and charged amino acids, but its effect will be stronger when combined with Martini polarizable water[51] and the recently released parameters for Martini with explicit polar amino acids.[52]

Bonded interactions were obtained by mapping UA systems to CG resolution. In order to do so, a UA simulation of a single cofactor molecule was simulated using the GROMOS topologies for 100 ns in a system containing 133 decane molecules. In general, the conformations sampled will, to some extent, depend on the solvent used and it might not be possible to find a CG topology that optimally mimics the conformations sampled in all solvents. Here we choose to base the CG topology on a simulation in (apolar) decane, since the cofactors are hydrophobic and will spend most time in apolar environments.

UA parameters obtained in the first step of the parametrization were used for the solute, while the model of decane was described earlier.[53] Distributions for all bonds, angles, and dihedrals were extracted from UA trajectories mapped to CG resolution, and from these distributions, initial guesses for the bonded parameters were obtained. Next, in an iterative procedure, the parameters were adjusted to obtain optimal distributions from a CG simulation while maintaining numerical stability. For every iteration, a 10 ns simulation of the CG cofactor molecule in 133 CG decane molecules was run, and the resulting distributions for all bonded interactions were obtained and compared to distributions from the mapped UA simulations. These steps were repeated until satisfactory distributions were obtained. As a synoptic measure of the similarity of two sets of parameters for a given molecule, we used the root-mean-square deviation (RMSD) of both the mean and the standard deviation of the distributions averaged over all bonds, angles, and proper and improper dihedrals.

Test Systems. The non-porphyrin-based cofactors (i.e., PQ9one, PQ9ol, UQ10one, UQ10ol, BCR) were simulated in a 128 lipid DPPC bilayer, using both CG and UA models. The CG simulations were started from a pre-equilibrated bilayer, obtained from www.cgmartini.nl. The bilayer consisted of 128 lipids solvated in 2000 CG water beads, corresponding to 8000

Table 1. Log P Values for the Partitioning of CG Cofactors between Water (W) and Octanol (Oct)[a]

	$\log_{10} P_{W\to Oct/Chx}$			
molecule	CG	UA	prediction	literature
chlorophyll A	$17.8^{1.9,b}$ $12.1^{0.2\,c}$	$27.6^{0.3}$		2.12^d
pheophytin	$21.7^{0.8\,b}$		10.4 ± 2.3	
heme b	8.9^b			0.95^d
β-carotene	$17.7^{0.2}$	$22.3^{0.2}$	6 ± 1	$17.62,^e$ 2.88^f
PQone$_9$	$20.2^{0.3,g}$ $27.9^{0.2\,h}$	$26.7^{0.4\,g}$		
PQone$_1$	$5.1^{0.1,g}$ $4.7^{0.1\,h}$	$4.9^{0.1,g}$ $3.7^{0.1\,h}$		$>3,^{g,i}$ $>3^{h,i}$
PQol$_1$	$2.6^{0.1,g}$ $0.6^{0.1\,h}$	$5.6^{0.1,g}$ $2.9^{0.7\,h}$		$>3,^{g,i}$ $0.23^{h,i}$
UQone$_1$	$3.8,^g$ 3.0^h	$6.1^{0.1,g}$ $5.4^{0.1\,h}$		$>3,^{g,i}$ $>3^{h,i}$
UQol$_1$	$2.8,^g$ 1.5^h	$6.5^{0.2,g}$ $5.9^{0.1\,h}$		$>3,^{g,i}$ $1.6^{h,i}$

[a]For CG plastoquinone (PQone), plastoquinol (PQol), ubiquinone (UQone), and ubiquinol (UQol), water−cyclohexane (Chx) partitioning is also reported. Subscript numbers attached to PQ/UQ one/ol indicate the tail lengths of the molecules: either 1 or 9 prenyl units. Statistical errors are given as superscripts. The average of log P's (prediction) obtained with online predictors[59−62] is indicated for a few molecules (see the Supporting Information, Table SI-2, for the full set of values). [b]Eight-bead ring topology. [c]Four-bead ring topology. [d]Reference 66. [e]Reference 63. [f]Reference 64. [g]W−Oct partitioning. [h]W−Chx partitioning. [i]Reference 65.

water molecules. For each of the cofactors, four molecules were placed in the water phase. The systems were subsequently energy minimized for 500 steps using the steepest descent algorithm and simulated for 10 000 steps at a small time step (5 fs) to properly relax possible overlap between water and cofactors. Starting configurations for the UA systems were obtained from a CG structure in which the cofactors had already moved into the bilayer, using the reverse transformation tool.[38] The Berger lipid parameters[54] were used to model the bilayer, together with SPC solvent. Simulations were performed using the setup described above for CG and UA systems, with the reference temperature set to 324 K (10° above the phase transition temperature of DPPC) and the pressure coupling scheme switched to semi-isotropic.

HEM was simulated bound to the protein bacterioferritin. The structure was taken from PDB entry 3E2C and converted to a CG model using the *martinize.py* script.[52] The Martini parameters for proteins[55] are used in combination with the ElNeDyn approach.[56] The interactions between the proteins and the HEM are purely nonbonded; e.g., no bonds are used to keep the HEM in place. The system was solvated by 5007 CG water beads (equivalent to 20 028 water molecules) and neutralized by adding 22 sodium ions. It was energy minimized for 500 steps using the steepest descent algorithm and simulated for 100 ps with position restraints on the protein backbone and a 10 fs time step. Finally, the system was simulated for 4 μs, using the settings as described above for Martini simulations.

Analysis. The conformations sampled by cofactors simulated in a bilayer system were analyzed using a cluster analysis algorithm described by Daura et al.[57] and implemented in the Gromacs tool g_cluster. An RMSD cutoff of 0.4 nm allowed the segregation of relevant conformations. Electron densities were calculated along the Z-axis of the box (coinciding with the membrane normal) using the Gromacs analysis tool g_density. The electrons of the atoms mapped to one CG bead were added together. To properly compare the densities obtained from UA and CG simulations, the UA trajectories were mapped to CG using the reverse transformation protocol[38] prior to the calculations of the electron densities. Diffusion rates were calculated from the mean square displacements (MSDs) averaged over similar molecules in the system. The diffusion in the x and y direction (D_x and D_y) is calculated separately to check for anisotropy. The linear regime of the MSD curve in

one dimension was fitted to MSD = $2 \cdot D \cdot t + c$, where D is the diffusion constant in m·s^{-1}, t is the time in s, and c is a constant accounting for initial fast diffusion. The final value is the mean of the D_x and D_y; the error is the standard error calculated over similar molecules.

■ RESULTS AND DISCUSSION

Parameterization of the UA Models. For most cofactor molecules, only modest changes were made with respect to the initial topologies generated by the PRODRG-server[40] or the Automated Topology Builder.[41] Adjustments of the bonded terms were needed to obtain a set of topologies consistent with each other. Final topologies were stable in simulations using an integration time step of 2 fs, with all bonds constrained. The full topologies in Gromacs format are provided in the Supporting Information or can be downloaded from www.cgmartini.nl. Below we detail the parametrization strategy for each of the cofactors individually.

Chlorophyll A and Pheophytin. The bonded and non-bonded interactions of both molecules are identical, except for atoms bound to the central magnesium absent in PHO. Both topologies contain three explicit hydrogens connected to the porphyrin ring (pointing out of the ring from groups 2/4/8 in Figure 1A). These hydrogens carry a +0.1 e charge and no LJ interactions. The magnesium ion in CLA carries a +1.0 e charge, which is compensated for by negative partial charges in the porphyrin ring. In PHO, the magnesium is replaced by two hydrogen atoms (bound to two opposing N atoms and pointing toward the center of the ring), each with a +0.1 e partial charge. Partial charges in the porphyrin ring are lower in PHO compared to CLA. Overall, both molecules are neutral.

The water/octanol partitioning coefficient of our CLA model is extremely far from the experimental value, 27 vs 2.12 (Table 1). To investigate the possible reasons for the difference, we predicted the water/octanol partitioning for PHO. There are no experimental values available, but as an alternative, we used four online predictors.[59−62] (PHO was chosen over CLA, since the Mg ion in CLA cannot be modeled by the commonly available predictors.) The predicted values show a reasonable spread (7.21 to 12.08 with a mean value of 10.4 ± 2.3, see the Supporting Information), but they all are considerably larger than the experimental value. The absence of the central ion in PHO cannot account for the difference with the experimental log P for CLA. It is also important to keep in mind that it is

likely that the predictor algorithms for PHO are not representative of CLA. Moreover, both the ring conjugation and the magnesium center are challenging features to account for making the UA model and the predictions quite hazardous. However, the results suggest that the experimental log P value for CLA might not be exact. See also the discussion below concerning the CG model.

Heme b. United-atom parameters for heme were already available in the GROMOS force field.[30] However, in order to make the topology fully compatible with the topologies of the other molecules, some of the bonded terms were refreshed. The nonbonded parameters were left unchanged.

β-Carotene. The topology for BCR was created using the ATB and then refined on the basis of the existing topology for retinol.[30] However, explicit hydrogens were added to double bonds in the conjugated chain. BCR is a symmetric molecule, basically consisting of two retinol molecules bound together by a double carbon−carbon bond after removal of the terminal alcohol groups.

Table 1 gives the partition coefficient, log P, for UA BCR. The corresponding values found in the literature diverge significantly. Cooper et al.[63] predicted a value of 17.62, while Liao et al.[64] reported a value of 2.88 based on UV measurements. Moreover, a series of online log P predictors[59−62] led to the average value of 6 ± 1 (see the Supporting Information for details). Our UA model gave a log P value of 22.3, which is larger than either of the values found in the literature and suggests that our model is too hydrophobic. The comparison of the atomistic partial charges of our model with those of other models revealed that our model might underestimate the polarity of the C double bond in the linker. This might contribute to the high hydrophobicity of our model. To test the effect of the partial charges on the BCR log P, we used another set of charges[29] and repeated the calculation. We found a log P value of 20.4, indicating that the partial charges are not the main contributor to the high hydrophobicity of BCR. The linker itself might be of importance, as one can see by comparing the log P of PQ1one to PQ9one (see below).

Plastoquinone and Plastoquinol. Plastoquinone and plastoquinol are *in vivo* typically found with nine prenyl units in the tail. However, the repeated parameters per prenyl unit make it easy to shorten or prolong the tail. In order to compare to partitioning experiments and quantum mechanical energy profiles, we also created models with a single prenyl-unit tail (PQ1one and PQ1ol) and an ethyl group instead of a prenyl tail (PQet).

The bonded and nonbonded interactions for both PQ9one and PQ9ol are equal, except for the ester or alcohol substituents of the ring. The distribution of the partial charges was taken from the charge calculations (see the Methods section) for both molecules separately. The tails are made of prenyl repeats. The hydrogen connected to the carbon involved in the prenyl double bond is explicitly represented, and carries a +0.185 e partial charge to mimic the strong polarization of the double bond. In addition, these hydrogens carry a small repulsive LJ term to avoid clashes with other atoms.

We calculated the potential energy of the dihedral angle describing connection of the PQ9one headgroup with the first prenyl tail unit (Figure 2). Zhang et al.[29] and Himo et al.[58] calculated this potential energy in a vacuum using a quantum mechanical model for plastoquinone and semiplastoquinone radical anion (PQet⁻, where the first prenyl tail unit is replaced by a radical anion ethyl group), respectively. When comparing
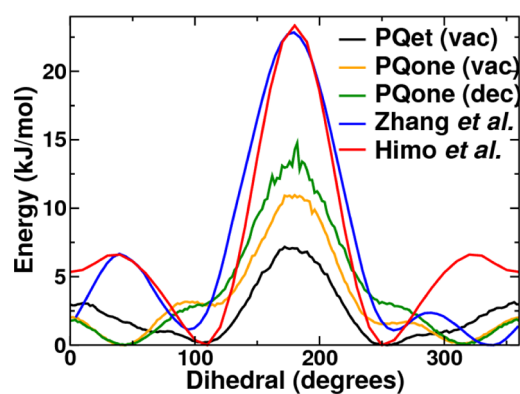

Figure 2. Energy profile of the dihedral angle modeling the connection between the plastoquinone head and tail. The energy distribution has been calculated for PQet (black line) where the polyprenyl tail has been replaced by an ethyl group, PQ9one in vacuum (yellow line) and decane (green line). The calculated distributions are compared to profiles from the literature for the same dihedral for similar systems: QM profile for PQ1one in a vacuum obtained by Zhang et al.[29] (blue line); QM profile of the PQet anion radical in a vacuum obtained by Himo et al.[58] (red line).

the energy profile obtained here for our model of PQet (neutral) to the one obtained by Himo et al. for PQet⁻, it can be seen that the shape is similar to the correct global maximum, although the energy barriers are reduced (Figure 2). The positions of the secondary maxima are slightly off at 10 and 350° (against 50 and 310° by Himo et al.) and too low. The distributions obtained for PQ9one in a vacuum reproduce the main features of the distribution found by Zhang et al.; however, the height of all barriers is off by approximately a factor of 2 and the positions of the secondary minima are shifted. When simulated in decane, the global maximum slightly increases, whereas most other features are similar. Although for both molecules the energy profiles do no perfectly match the QM energy profiles, the overall features are sufficiently well reproduced to give a proper united atom topology.

We determined the partition free energies for PQ9one, PQ1one/ol between water and octanol or cyclohexane (Table 1). Experimentally, the water/cyclohexane partition free energy has been determined for PQ1ol, while only lower bound values are available for PQ1ol water−octanol partitioning, and PQ1one water−cyclohexane and water−octanol.[65] The current model only partially reproduces these values. The absolute values for both PQ1one and PQ1ol are close to the experimental values, and the trends going from water/octanol to water/cyclohexane are correct. However, PQ1ol has a higher water/octanol partition coefficient than PQ1one, although one would expect the opposite.

Since partition free energies can be particularly sensitive to the atomic partial charges, we recalculated the charge distributions for PQone/ol using a higher level of theory and recalculated their partition free energies using the partial charges previously derived by others[29] (see Table 1 and the Supporting Information). The similarity of the results from both tests led us to the conclusion that for the present molecules the level of theory did not affect the charge distributions significantly. Their comparison with other sets confirmed that observation (see the Supporting Information). Furthermore, the details of the charge distribution do not significantly affect their partitioning behavior. It is thus not clear
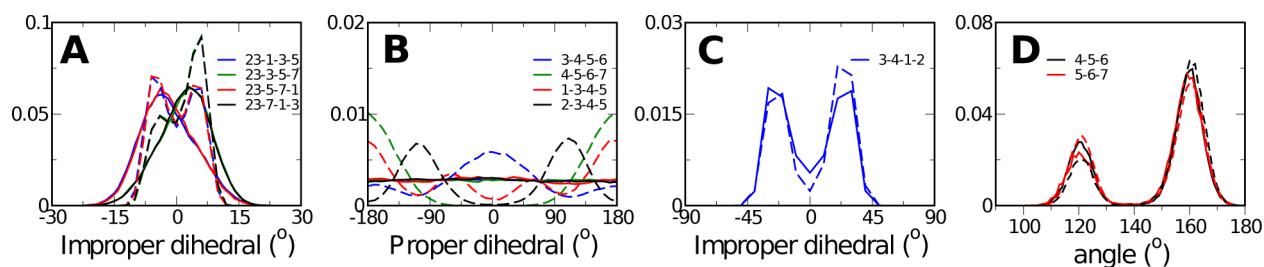
Figure 3. Probability distribution of selected dihedral angles observed in the CG (solid lines) and UA (dashed lines) simulations of chlorophyll (panel A) and β-carotene (panels B−D). The UA simulations have been mapped to CG to allow a direct comparison. The numbers in the legend correspond to the beads defining the dihedral angle (see Figure 1). Panel A shows improper dihedrals involving the metal center. Notice that due to the 4-fold symmetry of CLA all four lines overlap. Panels B and C show proper and improper dihedrals, respectively. Panel D illustrates the bimodal angle of the ring linker.

at this point why our model is not reproducing the PQone/ol difference. The models of solvent used might play a role.

Ubiquinone and Ubiquinol. Ubiquinone is *in vivo* typically found with a prenyl tail that is 10 prenyl units long. The bonded and nonbonded interactions for UQ10one and UQ10ol are similar, except for the ester or alcohol substituents of the ring, and very similar to those of plastoquinone and plastoquinol. The partial charges for the headgroup were calculated separately for UQ10one and UQ10ol. The parameters for the prenyl tail repeats are the same as those obtained for PQ9one and PQ9ol.

The partitioning behavior of our UQone/ol models is similar to the one described above for PQone/ol. The values have the correct order of magnitude when compared to experiments (Table 1) and respect the octanol/cyclohexane order. However, the expected higher preference of UQol for water over octanol than UQone is not reproduced. As in the case of PQ, a higher level of theory did not affect the atomic partial charges (data not shown) and the quality of the solvent (dry versus hydrated octanol, see the Supporting Information) did not have a major effect on these observations.

Parameterization of CG Models. The CG topologies presented in the following result in stable simulations with a time step up to 20 fs. Full Martini topologies in Gromacs format are provided in the Supporting Information or can be downloaded from www.cgmartini.nl. The mapping and typography of the cofactors are shown in Figure 1.

Chlorophyll A and Pheophytin. Topologies using four, six, or eight beads were tested to model the conjugated ring structure at the CG level. The six-bead variant was quickly discarded because it does not respect the 4-fold symmetry of the ring in contrast to the four- and eight-bead ring variants. Figure 1A shows the eight-bead mapping scheme. The water−octanol partition coefficients of the CLA topologies with a four- and eight-bead ring were calculated (Table 1). The log P values (12.1 and 17.8, respectively) indicate, as for PHO (21.7, eight-bead ring), that the topologies are much too hydrophobic when compared to the experimental value for CLA (log P = 2.12).[66] As discussed above for the UA model of CLA, the systematic large differences observed between the partition free energies we obtained raised enough doubt on the experimental values to settle with the current CG topologies. We favor the eight-bead ring variant, since it describes a thinner ring due to the use of S-beads, thus better reflecting the flat nature of the porphyrin ring.

This topology consists of 23 beads for chlorophyll A and 22 for pheophytin, mainly S-type particles in the ring and normal particles in the tail. In order to optimally maintain the

symmetry of the molecule, eight atoms have been doubly mapped: they are used in the determination of the center of mass (COM) of two adjacent beads. The masses of these atoms were halved for the COM calculation. The ring structure is very rigid and closely packed, giving rise to a high amount of stress. This stress was partially removed by defining exclusions (from the nonbonded interactions) between all of the ring beads. The central magnesium is modeled by a single SQ0 bead and has a +1.0 e charge, equal to the charge on the Mg in the UA topology. The four beads surrounding Mg carry a −0.25 e charge, to neutralize the structure. The charges on the Mg and surrounding beads mimic the strong polar nature of this part of the molecule. Placing the Mg at the last position of the topology allows generating pheophytin's topology by simply removing the Mg and the surrounding partial charges.

All CG bonded interactions were obtained by adjusting the parameters to give a satisfying reproduction of the distributions obtained from the UA simulations (see the Methods section). The CG topologies can be found in the Supporting Information. Over 37 bonds, 24 angles, and 7 proper and 15 improper dihedrals, the mean and standard deviation RMSDs between the CG and UA models were 0.004 ± 0.004 nm, 3.5 ± 1.3°, 9 ± 24°, and 5 ± 4°, respectively. These values are relatively high for the proper and improper dihedrals. For the proper dihedrals, this is due to three dihedrals in the tail that have flat distributions, and thus poorly defined means, for both the mapped UA and CG models. The four deviating improper dihedrals all involve the central magnesium, which visits two distinct states in the mapped UA model (above and below the plane of the ring), while it is normally distributed in simulations with the CG model, giving rise to different distributions (see Figure 3).

Heme b. The ring structure of HEM is very comparable to that of chlorophyll A. Therefore, a similar topology was used for both molecules, using eight beads to model the porphyrin ring (Figures 1C and 2C). Heme is modeled to carry two negative charges, as both propionate side chains (pK_a = 5.0) are deprotonated at physiological pH ∼ 7.0. Like for chlorophyll A, the CG topology of HEM models a compound that is too hydrophobic compared to experiment (log P_{CG} = 8.9 versus log P_{exp} = 0.95) (Table 1).[66] Notably, HEM is more hydrophilic than CLA, and the ratio of their CG log P values closely matches the ratio of the experimental values (log P_{HEM}/log P_{CLA} ≈ 2.4).

The topology of HEM contains 19 beads, of which all but two are S-type beads (Figure 1). In HEM, double mapping and bead exclusions are applied, as is the case for CLA keeping the correct symmetry and partially relaxing strain in the ring.

As for CLA, CG bonded interactions of HEM were derived by adjusting the bonded parameters to reproduce the distributions obtained from the UA simulations. Over 32 bonds, 24 angles, and 4 proper and 12 improper dihedrals, the mean and standard deviation RMSDs between the CG and UA models were 0.002 ± 0.002 nm, $1.7 \pm 1.8°$, $7 \pm 6°$, and $0 \pm 4°$, respectively.

β-Carotene. The BCR molecule is modeled by 10 beads (Figure 1C). The two rings and its substituents are modeled by three beads, of which two are S-type. The linker between the beads is divided into five equal parts, each modeled by a C4 bead with a 5:1 mapping. Nonbonded interactions between all beads in the ring and the first bead in the linker are excluded.

The partition coefficient for the CG model of BCR ($\log P = 17.7$) is close to the value found by Cooper et al. but high compared to most other reported values (Table 1). It is also significantly higher than predictions obtained from online predictors.

Over 11 bonds, 8 angles, and 7 proper and 2 improper dihedrals, the mean and standard deviation of the RMSDs were 0.007 ± 0.001 nm, $24 \pm 7°$, $6 \pm 24°$, and $11 \pm 3°$, respectively. Figure 3 shows the CG proper and improper dihedrals that do not give a well matching distribution. For symmetric pairs of dihedrals, only one is shown. The angles in the linker (4−5−6 and 5−6−7) show strong bimodal behavior and are rather extended. In order to correctly model the bimodal behavior, tabulated potentials were derived by combining two harmonic potentials with minima at 121 and 160°. The kinetics of the transition between the two conformations were reproduced by tuning the energy barrier in the CG potential. The resulting CG distributions match well with the UA distributions (see Figure 3D).

Due to the extended nature of the angles, applying a dihedral potential to the linker leads to numerical instabilities: if any of the bond angles contained in a dihedral angle are close to 180° (alignment of three consecutive beads), the definition of the dihedral angle becomes difficult and this gives rise to large forces leading to numerical instabilities. Solutions to this issue have been derived.[67] Preliminary tests indicated a potential improvement of the description of the dihedral angles. However, in what follows, we will only consider simulations using a topology without these dihedrals. The CG improper dihedrals (3−4−1−2 and 8−7−9−10) do deviate from mapped UA distributions; however, given the width of both distributions, the deviations are reasonable.

Plastoquinone and Plastoquinol. Both molecules have the same mapping (Figure 1D). The headgroup is represented by one C3 bead and two SNa beads for plastoquinone or one C3 and two SP2 beads in the case of plastoquinol. The tail is modeled by nine C3 beads, approximating a 1:5 mapping. The C3 bead was chosen on the basis of the water−octanol partitioning of *iso*-prenyl (-13 kJ mol^{-1}),[68] similar to the partitioning of C3 beads (-14 kJ mol^{-1}).[48] Note that the *iso*-prenyl unit differs from the repeated unit in *β*-carotene by one double bond, giving *iso*-prenyl a slightly more apolar nature.

The partition coefficients for the CG PQ1one and PQ1ol match the experimental values fairly well (Table 1). PQ1one has a log P higher than 3.0 for the partitioning between water and both apolar solvents. The PQ1ol water−cyclohexane partitioning is close to the experimental value ($\log P_{CG} = 0.56$ vs $\log P_{exp} = 0.23$). PQ1ol favors slightly too much water over octanol ($\log P_{CG} = 2.6$ vs $\log P_{exp} > 3.0$). The trends

between water and octanol/cyclohexane and between PQone/ol are correct.

Over 12 bonds, 10 angles, and 8 proper and 2 improper dihedrals, the mean and standard deviation RMSDs between the CG and UA models were 0.007 ± 0.007 nm, $6.1 \pm 5.1°$, $7.9 \pm 26.1°$, and $0.18 \pm 34.9°$, respectively. Note that to properly reproduce the bonded parameters the nonbonded interactions between the headgroup beads and the first tail bead are excluded.

Ubiquinone and Ubiquinol. Ubiquinone and ubiquinol have a very similar CG topology as plastoquinone/ol (Figure 1E). The ring is represented by three N0 beads and one SNa bead in the case of ubiquinone. The SNa and one of the N0 beads are switched to SP1 and Nda beads to model ubiquinol. This change reflects the more hydrophilic nature of an alcohol group compared to a carbonyl group. The tail, consisting of repeated *iso*-prenyl units, is identical to the ones of PQone/PQol and modeled by C3 beads. Water−cyclohexane and water−octanol partition free energies were calculated for both headgroups plus the first unit of the tail (Table 1). The CG models are consistent with the available experimental data.

As in the case of PQ, nonbonded interactions between the headgroup beads and the first bead of the tail are excluded. Over 14 bonds, 10 angles, and 3 improper dihedrals, the mean and standard deviation RMSDs between the CG and UA models were 0.003 ± 0.017 nm, $2.6 \pm 16.9°$, and $0.3 \pm 77.9°$, respectively.

Test System Simulations. In order to investigate the behavior of the CG cofactor topologies in a model of their physiological environment, we simulated HEM bound to a protein complex and BCR, PQ9one/ol, and UQ9one/ol in a lipid bilayer. The results were compared to the crystal structure of the protein complex and to analysis of UA simulations, respectively.

Heme b. Inside the PSII complex, HEM is bound in between the α and β subunits of cytochrome B559 or inside a pocket in cytochrome C550.[8] The central iron is coordinated by amino acid side chains of these proteins. To test if the CG HEM fits inside a protein binding pocket and the interactions with protein ligands are correctly modeled, we simulated it in a similar environment: between two bacterioferritin proteins. Bacterioferritin is a homodimer of two 158 amino acid proteins that together encapsulate one heme (PDB entry: 3E2C). The HEM iron is coordinated on both sides by a methionine. This system was chosen because the HEM is in between two independent protein domains so that it is possible that it escapes. It is thus a more stringent test than when HEM is buried in a protein, e.g., a cytochrome. Although this test does not guarantee a proper parametrization of the HEM, it does show that HEM can be successfully simulated while ligated to a protein. Note that in most applications one probably wants to completely prevent the disruption of the coordination of the HEM by adding a harmonic potential between the HEM iron CG bead and the coordinating molecules.

Over a 4.0 μs simulation, the complex remained stable and the heme stayed doubly coordinated. Omitting the first μs of the simulation as an equilibration period, the C_α RMSD of protein subunits A and B and the complex including the heme were 0.13 ± 0.01, 0.14 ± 0.02, and 0.23 ± 0.04 nm, respectively. In the (coarse grained) crystal structure, the distances between the HEM iron and the two coordinating methionines are 0.27 and 0.29 nm. The large van der Waals radius of the CG beads prevents these distances from
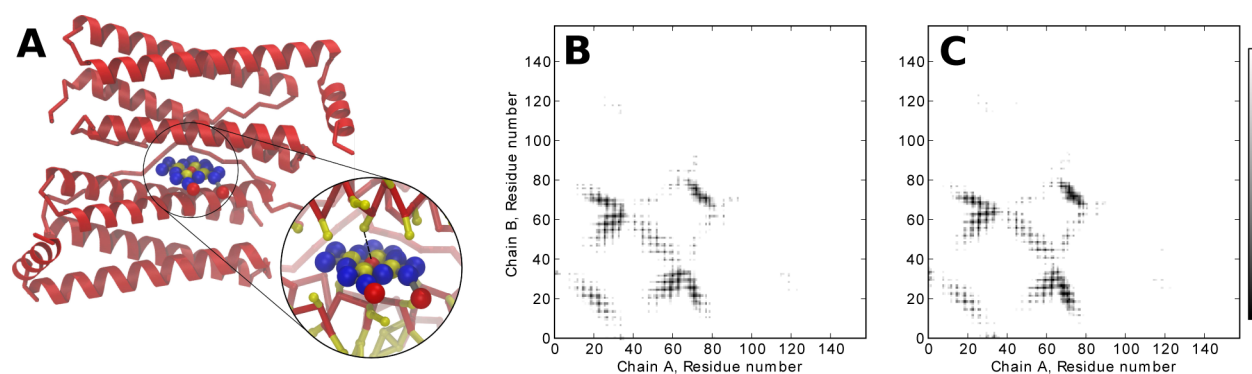
Figure 4. (A) Snapshot of a HEM molecule in between two bacterioferritin subunits after 1.5 μs of simulation. Proteins are shown in red; the HEM uses the same coloring scheme as in Figure 1. The inset gives a detailed view of the heme being coordinated by the methionine side chain. (B, C) Comparison of residue−residue contact maps between chains A and B in the crystal structure (B) and the mean over the last 500 ns of the simulation (C).
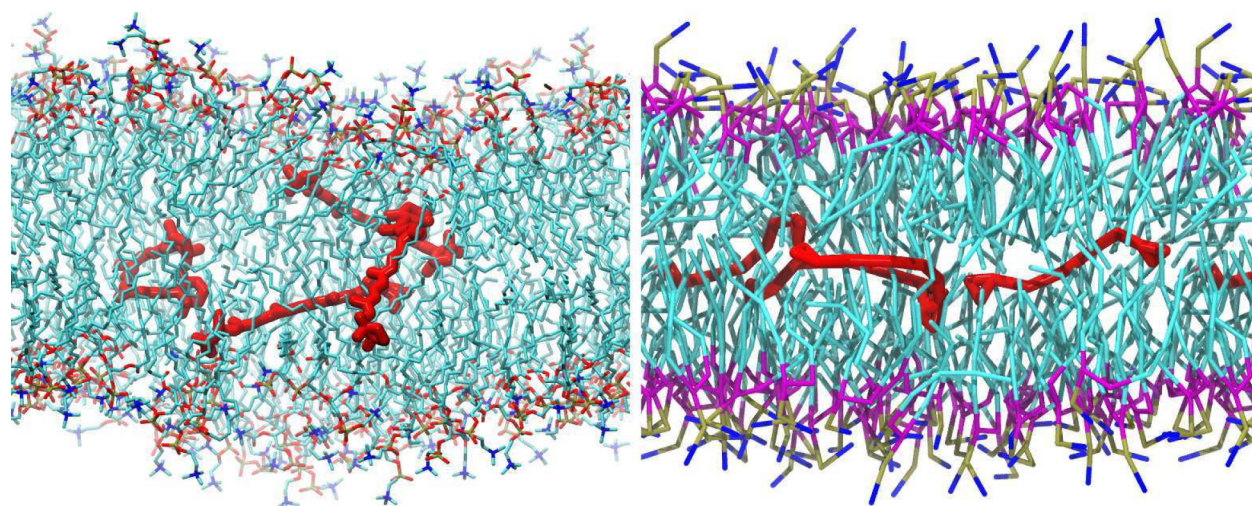
Figure 5. Snapshot of β-carotenes (red) at UA (left) and CG (right) resolutions in the midplane of a DPPC lipid bilayer.
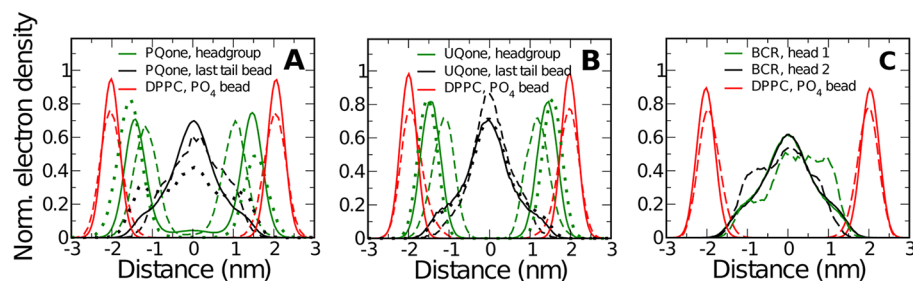
Figure 6. Normalized electron density plots for (A) PQ9one and PQ9ol, (B) UQ10one and UQ10ol, and (C) β-carotene in a DPPC bilayer. Densities obtained with the coarse grain model are compared to those obtained for the UA model mapped to coarse grain resolution. (A) Densities for the DPPC phosphate bead (red) are compared to those for the plastoquinone headgroup (green) and plastoquinone last tail bead (black). The solid lines are CG PQ9one, the dotted lines are PQ9ol, and the dashed lines are mapped UA PQ9one. (B) Densities for the DPPC phosphate bead (red) are compared to those for the ubiquinone headgroup (green) and ubiquinone last tail bead (black). The solid lines are CG UQ10one, the dotted lines are UQ10ol, and the dashed lines are mapped UA UQ10one. (C) Densities obtained for the DPPC phosphate bead (red) are compared to those for the CG (solid lines) and mapped UA (dashed lines) models of BCR. The profiles for BCR are averaged over four molecules and shown separately for the two headgroups of BCR (green and black lines). For CG, the profiles for both headgroups exactly overlap, but atomistically, the profiles differ, indicating that the simulation is not well converged. Phosphate bead densities from the different simulations are similar; only one copy is shown for clarity.

maintaining such small values during the simulation. Instead, the average distances over the simulations were 0.64 ± 0.06 and 0.51 ± 0.05 nm, which are close to the minimum distance between two CG beads. The addition of an explicit bond ($k = 20\,000$ kJ mol^{-1} nm^{-2}) between the iron and the methionines

allowed this deviation to be reduced to 0.33 nm. The residual deviation reflects the interaction of the iron and methionine side chain bead with other parts of the system. No further refinement was attempted on this system, but we have been able to obtain perfect agreement on other systems (data not

shown). The contact maps of the coarse grained crystal structure and the average over the last 3 μs of simulation (see Figure 4) show very similar patterns. Overall, the analysis showed that the complex is very stable. In a control experiment where all HEM bead types were set to (S)P5 (strongly hydrophilic), the HEM molecule left the cleft between both proteins. Upon leaving, the protein complex remains together; however, the RMSD does slightly go up (0.33 ± 0.02 nm averaged over the last 2 μs). In a second control experiment in which all beads were set to (S)C4 (strongly hydrophobic), the complex remained intact with a low RMSD value (0.21 ± 0.03 nm); however, the coordination of heme molecules was perturbed (the distance between HEM iron and methionine went up to 0.73 nm).

β-Carotene. In the PSII core complex, BCR protects nearby chlorophyll molecules against photodamage in high-light conditions.[69] Although it is mostly buried inside the protein complex, its behavior in a bilayer environment has previously been studied.[70] To test the quality of our CG model, we compared simulations of BCR in a DPPC bilayer at both CG and UA resolutions.

Four CG BCR molecules were initially placed in the water phase, and the system was simulated for 4 μs. One BCR entered the bilayer within 10 ns. The other molecules clustered in the aqueous phase and entered the bilayer together as a trimer after 70 ns. Inside the bilayer, the trimer broke down rapidly to never form again. BCRs adopted a straight conformation and spent most of their time at the bilayer midplane (Figure 5). To avoid a long and costly equilibration period, the UA starting structure was generated from the end conformation of the simulation at the CG resolution. The system was then simulated for 200 ns. The molecules mostly remained in an elongated conformation at the bilayer midplane and showed no specific clustering. The BCR molecules at both CG and UA resolutions reside approximately in the same region of the bilayer, as shown by the normalized headgroup electron density (Figure 6C). The slight difference between densities in both monolayers obtained for the UA model is indicative that the simulation has not yet reached equilibrium.

Due to their alignment with the membrane plane at the bilayer midplane, the BCR molecules are expected to diffuse faster than the surrounding DPPC lipids. The diffusion rates measured for BCR and DPPC at the CG resolution were (0.12 ± 0.01) and (0.08 ± 0.01) × 10^{-5} cm² s^{-1}, respectively. At the UA resolution, the diffusion rates were (0.03 ± 0.01) and (0.01 ± 0.01) × 10^{-5} cm² s^{-1}. At both resolutions, BCR diffuses faster than DPPC molecules, although the effect is slightly more pronounced at the CG level. Both BCR and the lipids diffuse faster at the CG resolution, by a factor of 4 and 8, respectively, of the same order of magnitude to the ones generally observed for the Martini force field.[71] Experimentally, no diffusion rate for BCR in a lipid bilayer is known. Jemioła-Rzemińska et al.[70] estimated the diffusion of BCR in a POPC bilayer to be 0.052 × 10^{-5} cm² s^{-1} from 4 ns UA simulations using the OPLS force field. This diffusion rate is slightly higher than the one obtained in this work for GROMOS, which could reflect the different types of lipids used (POPC vs DPPC).

Plastoquinone and Plastoquinol. The organization and dynamics of the CG model of PQ9one and PQ9ol in the thylakoid membrane (the "plastoquinone-pool") is of great interest in relation to the transport of electrons between different protein complexes in the photosynthetic pathway.[7] To test if our CG model properly models the behavior of PQ9one/

ol in a membrane environment, we simulated PQ9one/ol inside a DPPC bilayer and compared their location along the membrane normal, the conformations sampled, their flip-flop rate, and lateral diffusion to observations made at the UA resolution.

The systems containing four CG PQ9one/ol molecules were simulated for 4.0 μs each. In the analysis, the first 200 ns were omitted as equilibration. The cofactors were initially placed in the water phase but segregated into the lipid bilayer within tens of ns. At the UA resolution, a DPPC bilayer containing four embedded PQ9one molecules was simulated for 250 ns, of which the first 10 ns were discarded as an equilibration period. The UA starting structures were obtained by backmapping the final structures of the CG simulations (see the Methods section).

Figure 6A shows the electron densities of UA PQ9one and CG PQ9one/ol in a DPPC bilayer. In order to properly compare the CG and UA densities, the UA trajectory was mapped to CG before the analysis. For reference, the electron density distributions of the lipid phosphate beads are reported (red lines in Figure 6A) and their distances show that the thicknesses of the CG and UA bilayers are equal. At both resolutions, the head of PQ9one/ol (green lines) resides just below the headgroup region of the bilayer; however, the (mapped) UA headgroups are buried ~0.3 nm deeper into the bilayer. This difference indicates that the CG PQ9one headgroups are slightly more hydrophilic than the UA ones. The end of the tail of both UA and CG PQ9one resides mostly at the bilayer midplane.

To test the possibility to improve the behavior of PQ9one, we replaced the SNa headgroup particle type by SN0. This change gave slightly lower water−octanol and water−cyclo-hexane log P values but still larger than the lower bound given by experiments (log P_{exp} > 3). The modified headgroups were slightly (~0.1 nm) deeper buried, in better agreement with the UA model; however, the headgroups spent a considerable amount of the simulation time at the bilayer midplane, in contrast to the UA model. On the basis of this observation and the fact that the SN0 particle does not account for the hydrogen acceptor properties of the molecule (binding to hydrogen donors and positively charged molecules), we decided to settle with a SNa type particle.

The CG PQ9ol headgroup is on average found slightly closer to the bilayer interface than PQ9one, as one would expect from the more hydrophilic headgroup. The secondary maximum in the density distribution of the tail is due to the higher occurrence of tail back folding (see below). The four solutes in both the PQ9one and PQ9ol simulations are in contact with each other during a relatively large part of the time due to the high solute to lipid ratio (1:32 for both CG and UA). However, the contacts are mostly via the tails and there is no particular preference for the headgroups or tails to stack.

A cluster analysis of the molecular conformations sampled revealed three preferred orientations in both CG and UA resolutions: the PQ9one/ol head remains right below the lipid headgroup region of one monolayer, and (1) the tail forms an L-shape with the last part of the tail at the bilayer midplane (L conformation), (2) the tail extends to the opposite monolayer, sometimes with a small S-turn at the bilayer midplane (I conformation), or (3) the tail curls back up (back folding) to the headgroup (U conformation). Figure 7 shows examples of these conformations found for PQ9one at UA (left) and CG (right) resolutions.
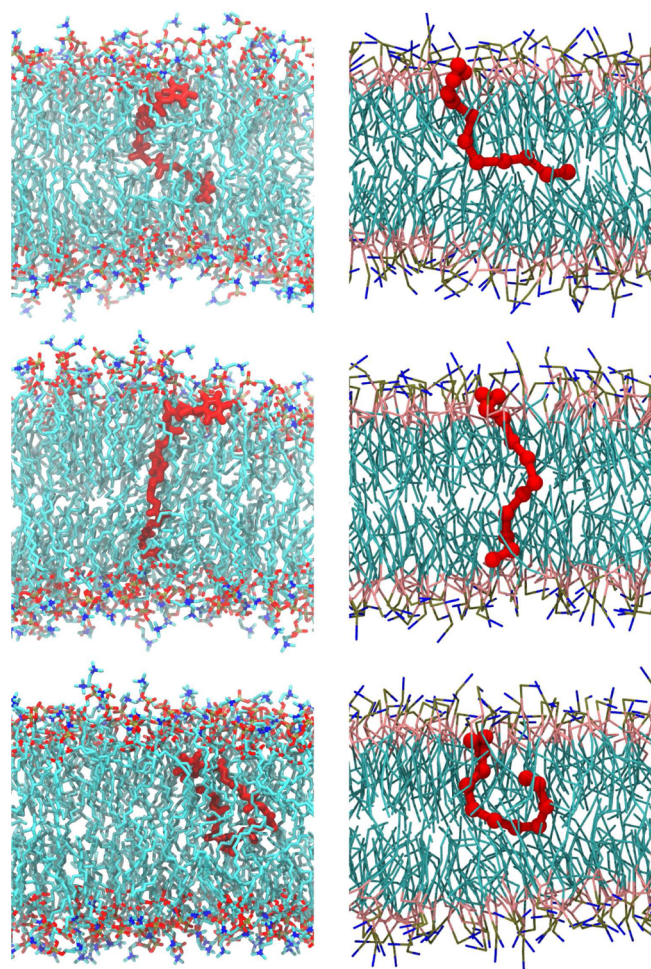
Figure 7. Side-by-side comparison of UA (left) and CG (right) plastoquinone in different conformations inside a DPPC lipid bilayer. Shown are the L conformation (top), the I conformation (middle), and the U conformation (bottom).

In the simulations containing four PQ9ones, the ratios L:I:U were 53(1):16(2):13(2)% and 46(14):18(3):13(3)% for the CG and UA resolutions, respectively. The values between brackets indicate the standard deviation calculated over the four molecules. The L, I, and U conformations account for 82 and 77% of the complete ensemble for the CG and UA resolutions, respectively. The remaining 18 and 23% consist of less well-defined, often less extended, conformations. For CG PQ9ol, the L:I:U ratio was 47(1):10(1):19(1)%, a significant shift from the I to U conformation. This stronger tendency of the tail to fold back toward the lipid headgroup is consistent with the electron density of the last tail bead of PQ9ol (Figure 6: the central density is more spread out) and is most likely due to the positioning of the more hydrophilic headgroup toward the water phase.

During the simulation, the PQ9one headgroup was observed to flip from one monolayer to the other, a so-called flip-flop event. In the CG simulation, 164 flip-flop events were observed, giving a flip-flop rate of 1.1×10^7 s^{-1}. In the UA simulation, the headgroup of one PQ9one molecule was observed to flip-flop twice. Although two events are insufficient to obtain a statistically reliable number, it indicates a flip-flop rate in the order of 10^6 s^{-1}. This means the flip-flop rate in the CG model is about 1 order of magnitude larger than in the UA model, in

line with results obtained for cholesterol flip-flop in DPPC bilayers (10^4–10^5 s^{-1} UA, 10^5–10^6 s^{-1} CG).[72] For CG PQ9ol, eight flip-flops were observed, a flip-flop rate of 0.05×10^7 s^{-1}. This much lower flip-flop rate as compared to PQ9one again reflects the more hydrophilic headgroup of PQ9ol.

Ubiquinone and Ubiquinol. The test simulations and analysis done for UQone and UQol were the same as those for PQ9one and PQ9ol. A tail of 10 *iso*-prenyl units was used for both the oxidized and reduced forms of UQ, since this is the most common species found in nature.[73]

In the simulations using a CG resolution, the UQ molecules quickly migrated into the bilayer and remained there for the entire simulation. Figure 6B shows the electron densities of UA UQ10one and CG UQ10one/ol in a DPPC bilayer. At the CG resolution, the head of UQ10one/ol (green lines) resides just below the headgroup region of the bilayer. At the UA resolution, however, the headgroups are considerably more buried toward the center of the bilayer. The slight asymmetry in the upper and lower monolayer lipid headgroup densities in the UA model is due to the asymmetric distribution of UQ10one molecules and indicates that the system is not fully equilibrated.

The three favored conformations described above for PQ9one/ol are also found for UQ10one/ol. In each system, a small portion of conformations could not be unambiguously assigned to either of the three conformations. Those molecules were left out of the analysis. In the simulations containing four UQol molecules, the L:I:U ratios were 62(1):12(2):15(1)% and 57(11):12(5):19(5)% for the CG and UA resolutions, respectively. In the case of UQone, the L:I:U ratios at the CG and UA resolutions were 51(2):11(1):17(2)% and 44(5):14(1):19(2)%, respectively. The slight preference for the U conformation compared to PQone/ol is due to the longer tail of UQone/ol. UQ10one and UQ10ol were not observed to flip-flop at either the UA or CG resolution. This can be explained by the increased hydrophilicity of the UQone/ol headgroups as compared to PQone/ol.

■ CONCLUSIONS

Topologies at united atom (UA) and coarse grain (CG) resolutions have been derived for cofactors associated with the PSII core complex found in the thylakoid membrane of the chloroplasts in plants, namely, plastoquinone and plastoquinol, heme b, chlorophyll A, pheophytin, and β-carotene. We also derived topologies for ubiquinone and ubiquinol, cofactors found in the mitochondrial membrane. The behavior of these molecules at both resolutions has been tested in different model systems. The CG topologies showed reasonable to good agreement with the UA ones and available experimental data.

At the CG resolution, irrespective of the number of beads used to model the ring of CLA and HEM, the molecules were found too hydrophobic when compared to experimental data, but our estimates were close to estimates from four online predictors. The eight-bead ring model was preferred over the four- and six-bead versions due to the respect of the 4-fold symmetry and the slightly flatter ring resulting from the use of S-type particles. While the bonded terms of CG CLA matched the (mapped) UA one, in HEM, two dihedrals showed dissimilar behavior without repercussion in the simulations: HEM remained bound in the crystal structure binding site of bacterioferritin.

BCR at the CG resolution perfectly reproduced the water−octanol partition free energy behavior described in the literature. Bonded interactions at the CG level showed good

agreement with the UA reference system. The dihedral angles in the elongated link between the two heads are difficult to model using conventional dihedral angle functions currently implemented in GROMACS. Nonetheless, the BCR at the CG resolution did properly mimic the behavior of the BCR inside a DPPC bilayer at the UA resolution.

Plastoquinone at the UA resolution shows reasonable agreement with experimental partition coefficients and qualitatively mimics energy profiles for the head tail dihedral obtained by QM calculations. At a CG resolution, plastoquinone and plastoquinol with different tail lengths show good agreement with experimental partitioning behavior and excellent agreement with UA bonded interactions. In a DPPC bilayer, CG PQ9one and UQ10one behaved similar to their UA counterparts.

The problems described above for the different topologies illustrate typical problems encountered when parametrizing molecules at CG resolution. Whereas the procedures to derive all-atom or united-atom topologies are better established[74] and may be automated,[41] the procedures to obtain CG topologies are less straightforward. The lack of experimental reference data, the constant trade-off between numerical stability and faithfully reproducing target properties, and the multitude of possible solutions make the parametrization process a great challenge with room for improvement.

The parametrized molecules represent some of the most important cofactors present in the thylakoid membrane. Together with recently published parameters for lipids found in the thylakoid membrane[75,76] and high resolution crystal structures of the thylakoid proteins,[8] these parameters will allow the simulation of the thylakoid membrane at both CG and UA resolutions and potentially using multiscale approaches.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information contains GROMOS and Martini topology files in Gromacs *.itp file format and the optional tabulated angle potential for the CG BCR topology in Gromacs readable.xvg file format. Additionally, the Supporting Information contains the details on the calculation of the partial charges for PQone and PQol at a higher level of theory and a recalculation of partition coefficients using a charge set obtained from the literature. It also contains data on the convergence of the calculation of log P in different solvents. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b00809.

AUTHOR INFORMATION

Corresponding Authors
*E-mail: X.Periole@rug.nl.
*E-mail: S.J.Marrink@rug.nl.

Present Addresses
§D.H.d.J.: Institute of Physical Chemistry, Corrensstr. 28/30, Muenster D-48149, Germany.
‖T.v.d.B.: Department of Physics and Astronomy and Institute for Lasers, Life and Biophotonics, Faculty of Sciences, De Boelelaan 1081, 1081 HV, Amsterdam, The Netherlands.

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Andersson, J. M. Photoregulation of the Composition, Function, and Structure of Thylakoid Membranes. *Annu. Rev. Plant Physiol.* **1986**, *37*, 93−136.

(2) Croce, R.; van Amerongen, H. Light-Harvesting in Photosystem I. *Photosynth. Res.* **2013**, *116*, 153−166.

(3) Scholes, G. D.; Fleming, G. R.; Olaya-Castro, A.; van Grondelle, R. Lessons from Nature About Solar Light Harvesting. *Nat. Chem.* **2013**, *3*, 763−774.

(4) Nelson, N.; Yocum, C. F. Structure and Function of Photosystems I and II. *Annu. Rev. Plant Biol.* **2006**, *57*, 521−565.

(5) Chaban, Y.; Boekema, E. J.; Dudkina, N. V. Structures of Mitochondrial Oxidative Phosphorylation Supercomplexes and Mechanisms for their Stabilisation. *Biochim. Biophys. Acta, Bioenerg.* **2013**, *1837*, 418−426.

(6) Mileykovskaya, E.; Dowhan, W. Cardiolipin-Dependent Formation of Mitochondrial Respiratory Supercomplexes. *Chem. Phys. Lipids* **2013**, *179*, 42−48.

(7) Guskov, A.; Kern, J.; Gabdulkhakov, A.; Broser, M.; Zouni, A.; Saenger, W. Cyanobacterial Photosystem II at 2.9-Å Resolution and the Role of Quinones, Lipids, Channels and Chloride. *Nat. Struct. Mol. Biol.* **2009**, *16*, 334−342.

(8) Umena, Y.; Kawakami, K.; Shen, J. R.; Kamiya, N. Crystal Structure of Oxygen-Evolving Photosystem II at a Resolution of 1.9 Å. *Nature* **2011**, *473*, 55−60.

(9) Ben-Shem, A.; Frolow, F.; Nelson, N. Crystal Structure of Plant Photosystem I. *Nature* **2003**, *426*, 630−635.

(10) Baradaran, R.; Berrisford, J. M.; Minhas, G. S.; Sazanov, L. A. Crystal Structure of the Entire Respiratory Complex I. *Nature* **2013**, *494*, 443−450.

(11) Sun, F.; Huo, X.; Zhai, Y.; Wang, A.; Xu, J.; Su, D.; Bartlam, M.; Rao, Z. Crystal Structure of Mitochondrial Respiratory Membrane Protein Complex II. *Cell* **2005**, *121*, 1043−1057.

(12) Lange, C.; Hunte, C. Crystal Structure of the Yeast Cytochrome *bc*₁ Complex With its Bound Substrate Cytochrome *c*. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2800−2805.

(13) Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yoshikawa, S. The Whole Structure of the 13-Subunit Oxidized Cytochrome *c* Oxidase at 2.8 Å. *Science* **1996**, *272*, 1136−44.

(14) Vassiliev, S.; Zaraiskaya, T.; Bruce, D. Molecular Dynamics Simulations Reveal Highly Permeable Oxygen Exit Channels Shared with Water Uptake Channels in Photosystem II. *Biochim. Biophys. Acta* **2013**, *1827*, 1148−1155.

(15) Palenčár, P.; Prudnikova, T.; Vácha, F.; Kutý, M. The Effects of Light-Induced Reduction of the Photosystem II Reaction Center. *J. Mol. Model.* **2009**, *15*, 923−933.

(16) Ogata, K.; Yuki, T.; Hatakeyama, M.; Uchida, W.; Nakamura, S. All-Atom Molecular Dynamics Simulation of Photosystem II Embedded in Thylakoid Membrane. *J. Am. Chem. Soc.* **2013**, *135*, 15670−15673.

(17) Sener, M.; Strumpfer, J.; Timney, J. A.; Freiberg, A.; Hunter, C. N.; Schulten, K. Photosynthetic Vesicle Architecture and Constraints on Efficient Energy Harvesting. *Biophys. J.* **2010**, *99*, 67−75.

(18) Postila, P. A.; Kaszuba, K.; Sarewicz, M.; Osyczka, A.; Vattulainen, I.; Róg, T. Key Role of Water in Proton Transfer at the Q₀-Site of the Cytochrome bc₁ Complex Predicted by atomistic Molecular Dynamics Simulations. *Biochim. Biophys. Acta, Bioenerg.* **2013**, *1827*, 761−768.

(19) Crofts, A. R.; Hong, S. J.; Wilson, C.; Burton, R.; Victoria, D.; Harrison, C.; Schulten, K. The Mechanism of Ubihydroquinone

Oxidation at the Q(o)-Site of the Cytochrome bc(1) Complex. *Biochim. Biophys. Acta, Bioenerg.* **2013**, *1827*, 1362−1377.

(20) Kaila, V. R. I.; Wikstrom, M.; Hummer, G. Electrostatics, Hydration and Proton Transfer Dynamics in the Membrane Domain of Respiratory Complex I. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 6988−6993.

(21) R.K.Le, B. J. H.; Iwuchukwu, I. J.; Bruce, B. D.; Cheng, X.; Qian, S.; Heller, W. T.; O'Neill, H.; Frymier, P. D. Analysis of the Solution Structure of Thermosynechococcus elongatus Photosystem I in n-Dodecyl-β-d-Maltoside Using Small-Angle Neutron Scattering and Molecular Dynamics Simulation. *Arch. Biochem. Biophys.* **2014**, *550*, 50−57.

(22) Arnarez, C.; Marrink, S. J.; Periole, X. Identification of Cardiolipin Binding Sites on Cytochrome c Oxidase at the Entrance of Proton Channels. *Sci. Rep.* **2013**, *3*, 1−8.

(23) Arnarez, C.; Mazat, J. P.; Elezgaray, J.; Marrink, S. J.; Periole, X. Evidence for Cardiolipin Binding Sites on the Membrane-Exposed Surface of the Cytochrome bc_1. *J. Am. Chem. Soc.* **2013**, *135*, 3112−3120.

(24) Karki, K.; Roccatano, D. Molecular Dynamics Simulation Study of Chlorophyll a in Different Organic Solvents. *J. Chem. Theory Comput.* **2011**, *7*, 1131−1140.

(25) Nilsson, J. A.; Eriksson, L.; Laaksonen, A. Molecular Dynamics Simulations of Plastoquinone in Solution. *Mol. Phys.* **2001**, *99*, 247−253.

(26) Autenrieth, F.; Tajkhorshid, E.; Baudry, J.; Luthey-Schulten, Z. Classical Force Field Parameters for the Heme Prosthetic Group of Cytochrome c. *J. Comput. Chem.* **2004**, *25*, 1613−1622.

(27) Palenčár, P.; Vácha, F.; Kutý, M. Force Field Development on Pigments of Photosystem 2 Reaction Centre. *Photosynthetica* **2005**, *43*, 417−420.

(28) Ceccarelli, M.; Procacci, P.; Marchi, M. An Ab Initio Force Field for the Cofactors of Bacterial Photosynthesis. *J. Comput. Chem.* **2003**, *24*, 129−142.

(29) Zhang, L.; Silva, D. A.; Yan, Y.; Huang, X. Force Field Development for Cofactors in the Photosystem II. *J. Comput. Chem.* **2012**, *33*, 1969−1980.

(30) Gunsteren, W. F. V.; Billeter, S.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; BIOMOS: Zürich, Switzerland, 1996.

(31) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. *Chem. Soc. Rev.* **2013**, *42*, 6801−6822.

(32) Wassenaar, T. A.; Ingólfsson, H. I.; Prieß, M.; Marrink, S. J.; Schäfer, L. V. Mixing MARTINI: Electrostatic Coupling in Hybrid Atomistic-Coarse-Grained Biomolecular Simulations. *J. Phys. Chem. B* **2013**, *117*, 3516−3530.

(33) Zavadlav, J.; Melo, M. N.; Marrink, S. J.; Praprotnik, M. Adaptive Resolution Simulation of an Atomistic Protein in MARTINI Water. *J. Chem. Phys.* **2014**, *140*, 054114.

(34) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.

(35) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A Generalized Reaction Field Method for Molecular Dynamics Simulations. *J. Chem. Phys.* **1995**, *102*, 5451−5459.

(36) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(37) Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116−122.

(38) Rzepiela, A. J.; Schäfer, L. V.; Goga, N.; Risselada, H. J.; de Vries, A. H.; Marrink, S. J. Reconstruction of Atomistic Details from Coarse-Grained Structures. *J. Comput. Chem.* **2010**, *31*, 1333−1343.

(39) Bennet, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245−268.

(40) Schüttelkopf, A. W.; Aalten, D. M. F. PRODRG: a Tool for High-Throughput Crystallography of Protein-Ligand Complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355−1363.

(41) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.* **2011**, *7*, 4026−4037.

(42) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.

(43) Kleywegt, G. J. Cristallographic Refinement of Ligand Complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *D63*, 94−100.

(44) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; et al. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

(45) Thole, B. T.; van Duijnen, P. T. A General Population Analysis Preserving the Dipole Moment. *Theor. Chim. Acta* **1983**, *63*, 209−221.

(46) Swart, M.; van Duijnen, P. T.; Snijders, J. G. A Charge Analysis Derived from an Atomic Multipole Expansion. *J. Comput. Chem.* **2001**, *22*, 79−88.

(47) Martyn, F.; Bush, I.; van Dam, H. J. J.; Sherwood, P.; Thomas, J. M. H.; van Lenthe, J. H.; Havenith, R. W. A.; Kendrick, J. The GAMESS-UK Electronic Structure Package: Algorithms, Developments and Applications. *Mol. Phys.* **2005**, *103*, 719−747.

(48) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.

(49) Marrink, S. J.; Fuhrmans, M.; Risselada, H. J.; Periole, X. In *Coarse Graining of Condensed Phase and Biomolecular Systems*; Voth, G. A., Ed.; CRC Press: Boca Raton, FL, 2009; Chapter The MARTINI Force Field.

(50) Periole, X.; Marrink, S. J. In *Methods inMolecular Biology*; Monticelli, L., Salonen, E., Eds.; Humana Press: New York, 2013; Chapter The Martini Coarse-Grained Force Field.

(51) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLoS Comput. Biol.* **2010**, *6*, e1000810.

(52) de Jong, D. H.; Singh, G.; Bennet, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 687−697.

(53) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Macedo, E. A.; Economou, I. G. 1-Octanol/Water Partition Coefficients of n-Alkanes from Molecular Simulations of Absolute Solvation Free Energies. *J. Chem. Theory Comput.* **2009**, *5*, 2436−2446.

(54) Berger, O.; Edholm, O.; Jähnig, F. Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature. *Biophys. J.* **1997**, *72*, 2002−2013.

(55) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.

(56) Periole, X.; Cavalli, M.; Marrink, S. J.; Ceruso, M. A. Combining an Elastic Network with a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531−2543.

(57) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.

(58) Himo, F.; Babcock, G. T.; Eriksson, L. A. Conformational Analysis of Quinone Anion Radicals in Photosystem II and Photosynthetic Bacteria. *J. Phys. Chem. A* **1999**, *103*, 3745−3749.

(59) Wang, R.; Gao, Y.; Lai, L. Calculating Partition Coefficient By Atom-Additive Method. *Perspect. Drug Discovery Des.* **2000**, *19*, 47−66.

(60) Wang, R.; Fu, Y.; Lai, L. A New Atom-Additive Method for Calculating Partition Coefficients. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 615−621.

(61) Tetko, I. V.; Bruneau, P. Application of ALOGPS to Predict 1-Octanol/Water Distribution Coefficients, LogP, and LogD, of AstraZeneca In-House Database. *J. Pharm. Sci.* **2004**, *93*, 3103−3110.

(62) Tetko, I. V.; Tanchuk, V. Y. Application of Associative Neural Networks for Prediction of Lipophilicity in ALOGPS 2.1 Program. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1136−1145.

(63) Cooper, D. A.; Webb, D. R.; Peters, J. C. Evaluation of the Potential for Olestra to Affect the Availability of Dietary Phytochemicals. *J. Nutr.* **1997**, *127*, 1699S−1709S.

(64) Liao, K.; Yin, M. Individual and Combined Antioxidant Effects of Seven Phenolic Agents in Human Erythrocyte Membrane Ghosts and Phosphatidylcholine Liposome Systems: Importance of the Partitioning Coefficient. *J. Agric. Food Chem.* **2000**, *48*, 2266−2270.

(65) Rich, P. R.; Harper, R. Partition Coefficients of Quinones and Hydroquinones and Their Relation to Biochemical Reactivity. *FEBS Lett.* **1990**, *269*, 139−144.

(66) de Vogel, J.; Jonker-Termont, D. S. M. L.; Katan, M. B.; van der Meer, R. Natural Chlorophyll but not Chlorophyllin Prevents Heme-Induced Cytotoxic and Hyperproliferative Effects in Rat Colon. *J. Agric. Food Chem.* **2005**, *135*, 1995−2000.

(67) Bulacu, M.; Goga, N.; Zhao, W.; Rossi, G.; Monticelli, L.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Angle Potentials for Coarse-Grained Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 3282−3292.

(68) Weast, R. *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, FL, 1986.

(69) Green, B. R.; Durnford, D. G. The Chlorophyll-Carotenoid Proteins of Oxygenic Photosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996**, *47*, 685−714.

(70) Jemioła-Rzemińska, M.; Pasenkiewicz-Gierula, M.; Strzałka, K. The Behaviour of β-Carotene in the Phosphatidylcholine Bilayer as Revealed by a Molecular Simulation Study. *Chem. Phys. Lipids* **2005**, *135*, 27−37.

(71) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004**, *108*, 750−760.

(72) Bennett, W. F. D.; MacCallum, J. L.; Hinner, M. J.; Marrink, S. J.; Tieleman, D. P. A Molecular View of Cholesterol Flip-Flop and Chemical Potential in Different Membrane Environments. *J. Am. Chem. Soc.* **2009**, *131*, 12714−12720.

(73) Metz, G.; Howard, K. P.; van Liemt, W. B. S.; Prestegard, J. H.; Lugtenburg, J.; Smith, S. O. NMR Studies of Ubiquinone Location in Oriented Model Membranes: Evidence for a Single Motionally-Averaged Population. *J. Am. Chem. Soc.* **1995**, *117*, 564−565.

(74) Lemkul, J. A.; Allen, W. J.; Bevan, D. R. Practical Considerations for Building GROMOS-Compatible Small-Molecule Topologies. *J. Chem. Inf. Model.* **2010**, *50*, 2221−2235.

(75) Lopez, C. A.; Sovova, Z.; van Eerden, F.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694−1708.

(76) van Eerden, F. J.; de Jong, D. H.; de Vries, A. H.; Wassenaar, T. A.; Marrink, S. J. Characterization of Thylakoid Lipid Membranes from Cyanobacteria and Higher Plants by Molecular Dynamics Simulations. *Biochim. Biophys. Acta, Biomem.* **2015**, *1848*, 1319−1330.