Erschienen in: Physical chemistry, chemical physics : PCCP ; 17 (2015), 34. - S. 22054-22063
https://dx.doi.org/10.1039/c5cp01140j

PCCP

PAPER

View Article Online
View Journal | View Issue

CrossMark click for updates

Cite this: Phys. Chem. Chem. Phys., 2015, 17, 22054

Derivation of coarse-grained simulation models of chlorophyll molecules in lipid bilayers for applications in light harvesting systems†

Ananya Debnath,‡*[a] Sabine Wiegand,[b] Harald Paulsen,[b] Kurt Kremer[a] and Christine Peter§[a]

The correct interplay of interactions between protein, pigment and lipid molecules is highly relevant for our understanding of the association behavior of the light harvesting complex (LHCII) of green plants. To cover the relevant time and length scales in this multicomponent system, a multi-scale simulation ansatz is employed that subsequently uses a classical all atomistic (AA) model to derive a suitable coarse grained (CG) model which can be backmapped into the AA resolution, aiming for a seamless conversion between two scales. Such an approach requires a faithful description of not only the protein and lipid components, but also the interaction functions for the indispensable pigment molecules, chlorophyll *b* and chlorophyll *a* (referred to as chl *b*/chl *a*). In this paper we develop a CG model for chl *b* and chl *a* in a dipalmitoylphosphatidyl choline (DPPC) bilayer system. The structural properties and the distribution behavior of chl within the lipid bilayer in the CG simulations are consistent with those of AA reference simulations. The non-bonded potentials are parameterized such that they fit to the thermodynamics based MARTINI force-field for the lipid bilayer and the protein. The CG simulation shows chl aggregation in the lipid bilayer which is supported by fluorescence quenching experiments. It is shown that the derived chl model is well suited for CG simulations of stable, structurally consistent, trimeric LHCII and can in the future be used to study their large scale aggregation behavior.

Received 25th February 2015,
Accepted 21st July 2015

DOI: 10.1039/c5cp01140j

www.rsc.org/pccp

Open Access Article. Published on 22 July 2015.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

1 Introduction

Chlorophyll *b* and *a* are two indispensable pigment molecules required for the assimilation of light energy in green plants. A significant percentage of the major light harvesting complex (LHCII) is comprised of these pigment molecules which, during photosynthesis, act as the antennae to capture light and transfer energy to the photosynthetic reaction center. The pigments are bound to the LHCII protein, a trimeric membrane protein that assembles with the other components of the photosystems in the thylakoid membrane of green plants. In addition to light harvesting, LHCII executes several photo-protective mechanisms in the case of excess light absorption.[1] A multitude of experimental techniques have been used to study LHCII.[1–3] The pigments' location and orientation are well resolved in the crystal structure of the protein:[4,5] they bind to the protein core at specific binding sites and the folding of the protein, *in vitro*, is accompanied by pigment binding.[2] Complementary to these experimental studies, molecular simulation can provide an atomic level picture of the interactions that are relevant to the pigment binding and the protein structure of LHCII. To study processes such as the trimerization of LHCII or the formation of larger LHCII aggregates, CG models are immensely useful, since atomic resolution will hardly be able to access the necessary time and length scales. In particular, in a multi-scale simulation setting where models on different levels of resolution are connected one can use the CG model to access the required length and time scales and then return to the AA level to study interactions locally including the chemical details.

In order to use such a multi-scale approach to investigate LHCII, one first needs model parameters for the components involved. While typical parameters for the protein and the lipid bilayer can be found in many classical force-fields, a challenging first task is to obtain a reliable model for the pigments, irrespective of the level of resolution. Recently, several mixed quantum/classical (QM/MM) models have emerged with chlorophylls (chl) in atomistic (AA) resolution. Here, properties such as

[a] Max Planck Institute for Polymer Research, 10 Ackermannweg, 55128 Mainz, Germany. E-mail: ananya@iitj.ac.in
[b] University of Mainz, Johannes-Von-Müller-Weg 6, 55128 Mainz, Germany

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5cp01140j

‡ Present address: Department of Chemistry, Indian Institute of Technology Jodhpur, Jodhpur 342011, India.
§ Present address: Universität Konstanz, Fachbereich Chemie, Fach 718, D-78457 Konstanz, Germany.

Konstanzer Online-Publikations-System (KOPS)
URL: http://nbn-resolving.de/urn:nbn:de:bsz:352-0-312162

 This journal is © the Owner Societies 2015

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

electronic transition,[6] electronic relaxation[7] or aggregation in different organic solvents[8] have been investigated. However, none of these models have been used to investigate long time biological processes which cannot be achieved solely by brute-force AA simulation. In recent years, a variety of CG methodologies for many biological systems have been developed.[9] Several CG models for proteins are built using knowledge based potentials, quasi-chemical approximations, elastic network models in combination with normal mode analysis.[10–12] Recently, cross-interactions of two independently derived CG models of lipids and peptides have been tuned to obtain a successful combination of parameters. Using these parameters the energetics for individual amino-acid side chain insertions into a lipid bilayer are reproduced well.[13] The MARTINI CG force-field for proteins and lipids is developed paying attention to the partition free energy between polar and apolar phases of a large number of chemical building blocks.[14–17] To accelerate the speed of the simulation in the presence of relevant chemical details, AA and CG resolutions can be coupled with either pre-defined sets of atoms or resolution exchange simulations across predefined interfaces.[18–20] These techniques are applied to binary mixtures, proteins through virtual CG sites and membranes.[21,22]

In our simulations of the pigmented LHCII, we develop and parameterize CG models for chl pigments with the aim of using them for CG and multi-scale simulations of pigmented LHCII complexes in lipid bilayer systems. An AA model is used to derive the CG model, and after the CG simulations, AA representations of the system can be retrieved *via* backmapping. The information transfer from the AA to the CG scale is direct and systematic for CG bonding potentials which are based on the structural properties of AA simulations, whereas the non-bonding potentials are parameterized and embedded into the existing MARTINI force-field. To start with, chl *b* and chl *a* are parameterized in the presence of the lipid bilayer. (Note that, the notation "chl *b*" or "chl *a*" specifically referred to one of the two chemically distinct chl molecules. For procedures that were equally applied to both types of pigments we will use the notation "chl" throughout the text.) This reference system for parameterization was chosen for the following reasons: the chl–lipid interactions are highly relevant for the formation and behavior of the LHCII protein–pigment complex in the lipid bilayer. *In vitro* studies have shown that the folding of the LHCII apoprotein and the pigment binding to the protein are tightly coupled processes. In the LHCII monomer, many chl pigments are situated in the outer region of the protein, effectively forming an interface between protein and lipids. Consequently, chl–lipid interactions are important for the assembly and stability of the pigmented protein trimer. Additionally, the chl–lipid system is also a more tractable reference system for which the interactions between the MARTINI standard force-field and chl can be compared to the fully pigmented LHCII membrane protein complex. To fit in the line of the general MARTINI parameterization philosophy which focuses on partitioning free energies in polar and apolar phases for different chemical building blocks, the non-bonded parameters of chl are chosen such that the partitioning of the CG chl beads between hydrophilic and hydrophobic regions in the bilayer is correctly represented compared to the AA reference. The bonded interactions in the CG pigments are derived such that the CG model reproduces the shape and the conformational behavior of the AA chl molecules. The degree of coarse graining for the CG beads is in general similar to MARTINI. Details on mapping and coarse graining strategies are given in the model development section. We carry out long-time scale simulations of large chl/lipid systems. We show that we can re-introduce atomistic details into the CG structures to start AA simulations in order to extract detailed chemical interactions of chl occurring on long time scales. We find chl aggregation in the bilayer, which is validated by comparison to fluorescence quenching experiments. Finally, we show a first application of our chl CG model to a pigmented LHCII protein trimer in a lipid bilayer. We aim to extend the applications to investigate larger aggregates of LHCII trimers in the future.

2 Atomistic simulations

2.1 Computational details

We first set up an AA reference simulation with two chl molecules (one chl *b* and one chl *a*) in the two separate leaflets of an already equilibrated, fully hydrated DPPC bilayer (denoted as system S1 in Table 1). The initial configurations of two different chl are taken from the crystal structure of LHCII (1RWT.pdb).[4] The energy minimized[23] configuration is equilibrated for 110 ns. Next a 100 ns *NPT* production run is carried out at 323 K using the velocity rescale thermostat[24] and at 1 bar using the Parrinello–Rahman semi-isotropic pressure coupling.[25] A time step of 1 fs is used with constraints on all bonds involving hydrogen. All water molecules are constrained using the SHAKE algorithm. All simulations have been performed using GROMACS-4[26–28] using twin range van der Waals cut-offs (1.0 nm and 1.4 nm) and a particle-mesh Ewald real-space cutoff of 1 nm with a 0.12 nm grid size. Kukol's force-field parameters are used for DPPC.[29]

Table 1 System compositions (number of DPPC lipids and chl molecules) and box parameters for the different chl/lipid simulations analyzed in this study. S1: AA system; S2M1C1, S2M2C1, and S2M3C1: CG systems with only chl *b* and parameter sets M1, M2 and M3, respectively; S2M1C2, S2M2C2, and S2M3C2: CG systems with only chl *a* and parameter sets M1, M2 and M3, respectively; S3: AA systems with tails removed from chl *b* and chl *a*; S4: CG systems with tails removed from chl *b* and chl *a*; having only the ring region excluding the tail. Also shown: membrane thickness as determined from the distance between the maxima in the lipid head group densities along the bilayer normal

System	Number of molecules			Box length, nm			Membrane thickness, nm
	Lipid	Chl *b*	Chl *a*	L_x	L_y	L_z	
S1	120	1	1	6.09	6.11	7.10	3.97
S2M1C1	480	8	0	12.28	12.31	5.51	4.07
S2M2C1	480	8	0	12.36	12.39	5.43	4.07
S2M3C1	480	8	0	12.28	12.31	5.51	4.04
S2M1C2	480	0	8	12.36	12.39	5.69	4.09
S2M2C2	480	0	8	12.38	12.41	5.68	4.03
S2M3C2	480	0	8	12.42	12.45	7.96	3.97
S3	120	1	1	6.07	6.09	7.09	3.90
S4	480	4	4	12.39	12.43	9.29	3.95

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

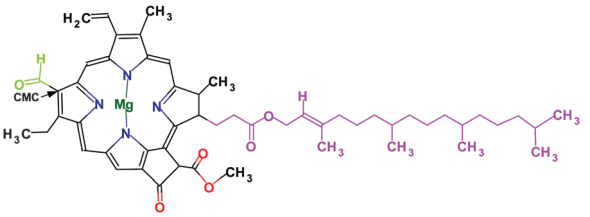

Fig. 1 Chemical structure of chl *b*. The aldehyde group presented in green color is replaced by a methyl group in chl *a*. The carbon atom closest to this aldehyde group is labeled as CMC in the discussion of the behavior of both chl *b* and chl *a* in the lipid bilayer. The atoms of the tail that are removed to study the behavior of the ring system are shown in pink.

We use the SPC/E water model.[30] For the chl molecules, GROMOS53a6 parameters are used for the rings (in analogy to the parameters for the heme group) and Kukol's parameters for the aliphatic tails. Partial charges for the Mg and N atoms are obtained by carrying out an *ab initio* calculation using the B3LYP/6-31G* basis set and by rescaling the partial charges in agreement with the typical partial charge distributions within the GROMOS force-field (final partial charges and atom types are shown in Table 1 of the ESI†).

To better distinguish between the influence of the ring and the aliphatic tail on the behavior of chl, we have also simulated a model system without tails (shown in pink color in Fig. 1; designated as S3 in Table 1).

2.2 Chl *b* and *a* in a DPPC bilayer

Snapshots of AA chl (the chemical structure is shown in Fig. 1) in the lipid bilayer are shown in Fig. 2 at different time steps. Fig. 2(a) shows the initial placement of both chl *b* and chl *a*, which corresponds to the typical placement of the chl around the LHCII.[4] The chl molecules do not remain in this initial region, instead both molecules move from the hydrophobic mid-plane of the bilayer towards the polar head groups. The polar groups in chl (shown in a red VDW representation) are attracted to the polar head groups of DPPC and the central Mg of the porphyrin ring is coordinated by a DPPC phosphate oxygen. On the other hand, the hydrophobic aromatic porphyrin ring favors to reside in the hydrophobic tail region of the DPPC bilayer. As a combined effect of polar–polar and hydrophobic interactions, the chl rings reside in the hydrophobic region of the lipid tails, but are inclined towards the polar head groups (see Fig. 2(b) and (c)). Moreover, due to the coordination between the central Mg and the DPPC phosphate, chl molecules pull some of the lipid head groups deeper into the hydrophobic region of the DPPC bilayer leading to a local disorder near the chl ring without distorting the overall membrane structure. In the following paragraphs, these AA simulations will provide reference data for the parameterization of the CG model.

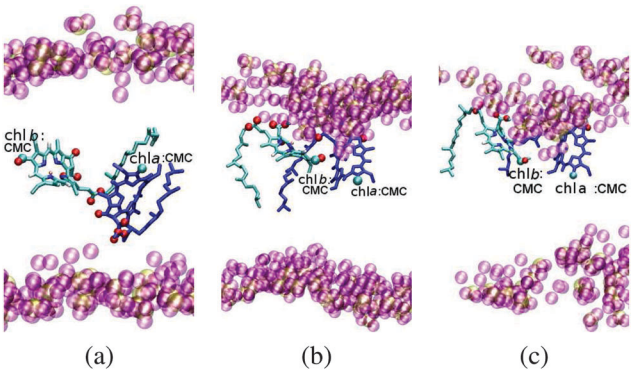

Fig. 2 Snapshots of AA chl *b* (cyan) and chl *a* (dark blue) molecules in the DPPC bilayer (a) at 0 ns, (b) after 100 ns, and (c) after 200 ns. The head group of the DPPC lipids is shown in (magenta and yellow) VDW representation. Chl *b* and chl *a* are shown in licorice representation with the polar oxygen atoms highlighted as red and the CMC carbons (see Fig. 1) as cyan spheres.

3 Developing the CG model

3.1 Mapping scheme

In the present mapping scheme, all AA beads are uniquely mapped onto CG beads in such a way that, the coordinates of the centers of the CG beads are defined by the centers of mass of the respective constituent atoms. In general, the degree of coarse graining follows the standard MARTINI approach for aromatic and aliphatic groups (the full mapping scheme is presented in Table 2 in the ESI†), with the exception of the central beads of chl: to be able to maintain the crucial and highly conserved atomic level interactions between the chl centers and specific amino acids of the protein (ESI,† Table 1 in ref. 4), the central Mg and N atoms are represented by a 1:1 (CG:AA) mapping with an assignment of partial charges similar to the AA level (see Table 3 in the ESI†). Chl *b* differs from chl *a* by the presence of one aldehyde group (shown in green in Fig. 1) which is represented by an explicit CG bead. For DPPC and the LHCII protein, the MARTINI[15] mapping is employed.

3.2 Bonded and non-bonded potentials

To derive the bonded potentials of CG chl, we assume that (i) the non-bonded and bonded interactions can be parameterized separately, and (ii) the different degrees of freedom which contribute to the bonded potentials are independent of each other.[31–35] Then, the conformational probability distribution, $P^{CG}(r, \theta, \phi, T)$, can be written as follows,

$$P^{CG}(r, \theta, \phi, T) = P^{CG}(r(T))P^{CG}(\theta(T))P^{CG}(\phi(T)), \quad (1)$$

where r, θ, ϕ, and T denote the bond length, angle, dihedral and temperature, respectively. Once the individual probability distributions are obtained from the canonical sampling of the AA reference state (S1), we invert these distributions to obtain the respective potentials of mean force, $U_q(q)$, as follows,

$$U_q(q) = -k_B T \ln(P_q^{CG}(q)) + \text{const}_q, \quad (2)$$

where $q = r, \theta, \phi$.

Once we obtain the bonded potentials, we fit them with the standard analytical interaction functions used in classical forcefields (in the present case the GROMOS forcefield). Our CG model for chl *b* has 33 CG bonds, 33 CG angles and 31 CG

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

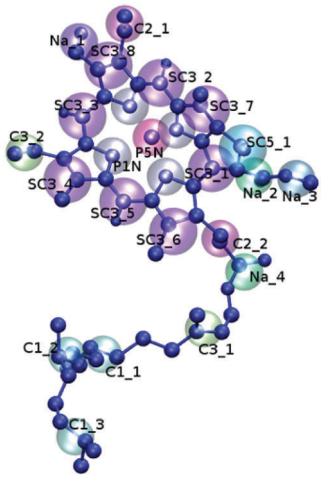

Fig. 3 Mapping of CG beads onto the chemical structure of the chl *b* molecule. Names of bead types indicate the use of standard MARTINI bead types.[15] The details of the mapping scheme are shown in Table 1 in the ESI.† The bead types shown for chl *b* are used for model M3. The two new bead types introduced in our CG model (corresponding to the central Mg and N atoms) are labeled as P5N and P1N, respectively. Chl *a* follows the same mapping scheme except for omitting the bead Na_1. The figure has been created using VMD.[48]

improper dihedrals, and chl *a* requires 32 CG bonds, 31 CG angles and 30 improper dihedrals, respectively.

The non-bonded potentials are parameterized to fit the MARTINI force-field.[15] We use standard MARTINI Lennard-Jones bead types for all the chl beads except for the beads in the center. Here, we introduce two new bead types P5N and P1N, for central Mg and N, respectively, by refining the interaction strength, ε_{ij} and the effective bead size, σ_{ij}, from the MARTINI parameters as will be described in more detail below. We have tested two bead types, SC4/3 for peripheral aromatic ring beads to reproduce the atomistic density distributions of the ring system within the bilayer. The other CG sites which are not directly part of the ring, but connected to it, are assigned to other bead types, shown in Fig. 3. The interactions between the newly introduced bead types with the remaining bead types have been calculated using geometric mixing which leads to a reasonable match between AA and CG non-bonded properties, such as, density distributions and orientational distributions.

3.3 Different CG models

Using typical sizes and interaction strengths from the MARTINI force-field for the central beads (Mg and N) of chl and then placing the chl in the protein environment of the LHCII result in an unstable simulation due to steric clashes. Apparently, the central beads of the chl molecule are of particular importance for the interaction with the protein. Overestimating their size by insisting on a typical CG mapping results in instabilities. The interactions between these central beads of chl and specific amino acids of the protein trimer in LHCII are found to be highly conserved in experimental studies (see, ref. 4 ESI,† Table 1). Similar interactions were found to be conserved in our AA simulations of the LHCII trimer (unpublished data), which indicates that

Table 2 Parameters (σ and ε of the central P5N (Mg) and P1N (N) bead types) for the three models M1, M2 and M3. The polarity of the porphyrin ring carbon is reduced from M1 to M2. The corresponding beads are denoted by the standard MARTINI[15] bead type SC3 in M2 instead of SC4 in M1. In M3, aldehyde, ketone and ester functions are represented by less polar bead type Na. Bonded potentials for M1, M2 and M3 are obtained by Boltzmann inverted distributions and further refinements, respectively

Model	σ_{P5N} (nm)	ε_{P5N} (kJ mol^{-1})	σ_{P1N} (nm)	ε_{P1N} (kJ mol^{-1})	Bead type ring carbon	Bead type aldehyde/ ketone/ester
M1	0.28	0.56	0.46	0.45	SC4	P2/3/4
M2	0.33	0.56	0.51	0.45	SC3	P2/3/4
M3	0.33	0.56	0.51	0.45	SC3	Na

such interactions play crucial roles in the stability of the trimer. In order to capture the correct interplay between the central beads and the protein in our CG model, we have consciously chosen a 1 : 1 mapping scheme from AA to CG resolution for the central beads so that our CG simulations of LHCII are stable and structurally consistent with the AA data as well as the X-ray crystallographic data.

Therefore, we fine tune the parameters for these beads based on two criteria, (1) the excluded volume (in the parameterization this is done with respect to the lipids) and (2) the preferential distribution within the lipid bilayer. To start with, we calculate radial distribution functions (RDFs) of the central bead, Mg and N (bead type P5N and P1N, respectively) and the atoms in the lipid tails (bead type C1). The position of the first maxima of the RDF gives an indirect estimate of the excluded volume in terms of an effective bead size. Fig. S2 in the ESI† shows that the σ_{ij} and ε_{ij} values from the AA and MARTINI show two extreme cases of too small or large excluded volume within which the values of σ_{ij} and ε_{ij} are refined. We varied the values of σ_{ij} and ε_{ij} starting from a typical AA Lennard Jones (LJ) interaction to a typical CG LJ interaction in a stepwise manner for both P5N and P1N. Only a small selection of parameter sets is presented here. We show three models (*i.e.* sets of σ_{ij} and ε_{ij} for P5N, P1N and the other chl specific bead types) which are from now on referred to as M1, M2 and M3 (in Table 2).

3.4 Computational details: CG simulations

The atomistic chl and DPPC molecules are mapped[36] to CG degrees of freedom in such a way that there is either one chl *b* or *a* present per leaflet in a 120 CG DPPC bilayer and then the system is replicated twice both in the *x*- and *y*-directions. The systems with chl *b* or *a* are simulated by the three CG models, M1, M2 and M3 and are referred to as S2M1C1, S2M2C1, S2M3C1 (for chl *b*) and S2M1C2, S2M2C2, S2M3C2 (for chl *a*) (Table 1), respectively. The energy minimized structures are equilibrated for 10 ns. A 1 μs long *NPT* production run is carried out at 323 K using the velocity rescaling thermostat[24] (τ_T = 0.5 ps) and at 1 bar using the semi-isotropic Parrinello–Rahman pressure coupling method (τ_P = 1.2 ps). The LJ and electrostatic potentials are shifted from 0.9 to 1.2 nm and from 0 to 1.2 nm, respectively. The neighbourlist is updated for every 5 steps. We use a time step of 2–4 fs which is limited by the stiffest potential in the system.

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

3.5 Backmapping

From the CG simulation trajectory, structures in AA resolution are generated *via* a backmapping procedure. To do that, we first re-introduce AA particles into the corresponding CG beads obtained from the systems S2M2C1 and S2M2C2 using model M2 (see Table 1) using an algorithm implemented in GROMACS.[37,38] After equilibration from 1.6 ns *NPT* runs, a 10 ns production run is carried out following similar conditions as in the initial AA runs. All the results from BM and AA simulations are reported after representing the AA coordinates into CG representation.

4 Results and discussion

4.1 Structural behavior

We have simulated the pigment/bilayer system using different CG models, reporting here about models M1, M2 and M3. First, we compute various intra-molecular distributions of distances between chl beads, which are not directly connected by bonds, angles or dihedrals. These distances are not introduced in the bonded parameterization and they are representative of the overall shape of the molecules. Thus, the effect of combining the parameterizations of bonded and non-bonded potentials can be assessed from these distributions. We first analyze the distributions across the porphyrin ring of chl (Fig. 4). Due to the refined bonded interactions, the distributions obtained from models M2 and M3 agree better with the AA distribution than those of model M1. Fig. 4 also shows the distribution of the end to end distance of the tail of chl. Note that the tail conformations are not exclusively a probe for the bonded interaction functions, since they are sensitive also to the spatial distribution and anchoring of the pigments in the lipid bilayer and therefore the non-bonded interactions with the lipids. Our CG models (M1, M2 and M3) cover the range of internal distance distributions reasonably well both for chl *b* and chl *a*. It has been shown earlier that the tail distribution of the AA chl *a* molecule is solvent dependent.[8] This supports the fact that the parameterization environment for chl should be the bilayer, since we use this CG chl model for the LHCII complex in the lipid bilayer. For the other bond and angle distributions involving not immediately connected beads across the porphyrin ring the match is comparable for M1, M2 and M3, both for chl *b* and chl *a* (Fig. S3 in the ESI,† shown for M1 and M2).

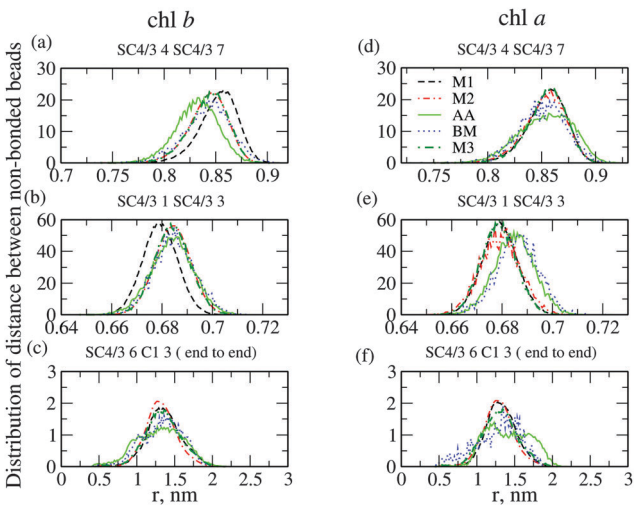

Fig. 4 Distance distributions between CG beads across the chl *b* (left panels) and chl *a* (right panels) rings and the end to end distributions for the tails. Green solid lines: AA; black dashed lines: CG M1; red dash dotted lines: CG M2; dark green dash dotted lines: CG M3; blue dotted lines: AA after backmapping from CG M2. Further data such as angle distributions are shown in the ESI.†

4.2 Partitioning of chl within the bilayer

We calculate the distributions of the positions of certain chl *b* and chl *a* atoms within the lipid bilayer (along the bilayer normal). This is an indirect measure of the partitioning behavior of different parts of the chl molecules in the DPPC bilayer. The distribution of the chl molecules along the bilayer normal is of particular importance since it reflects a partitioning free energy between the more hydrophobic membrane center and the more hydrophilic head-group region. This is a property that fits well with the general MARTINI parameterization philosophy which relies on transfer of free energies of reasonably small chemical compounds representing CG units between hydrophilic and hydrophobic media. Unfortunately a direct approach *via* transfer of free energies is unfeasible for the chl center since the pigment cannot be broken down into representative fragments. The number of density profiles is shown as a function of the distance from the mid-plane of the bilayer. Fig. 5(a) and (c) present the density distributions of the ring beads of chl *b* and chl *a*. As discussed in Section 2.2, the ring region of AA chl remains in the hydrophobic region of the lipids. In the CG model M1, the density distributions

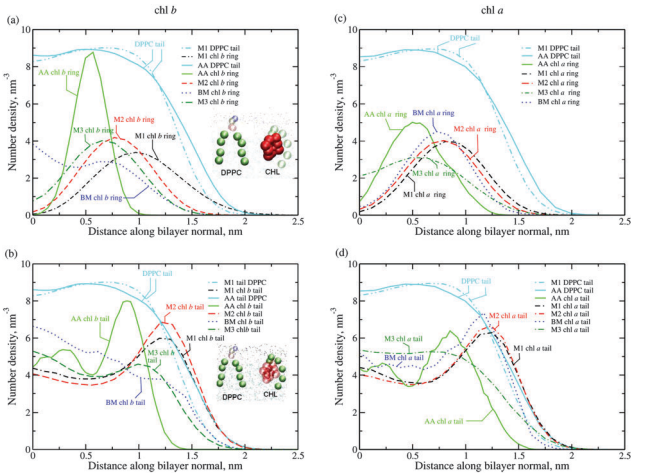

Fig. 5 Number density distributions of ring and tail beads of chl *b* (left panels) and chl *a* (right panels) as a function of the *z* position within the lipid bilayer (analyzed along the bilayer normal). Upper panels: distribution of chl ring beads (enhanced by a factor of 10 compared to the number density of the DPPC tail beads for clarity). Lower panels: distribution of the chl tail beads (enhanced by a factor of 50 compared to the number density of the DPPC tail beads for clarity). Cyan blue solid lines: DPPC tail beads, AA; cyan blue dash-dotted lines: DPPC tail beads, CG; green solid lines: chl, AA; black dot-dashed lines: chl, CG M1; red dashed lines: chl, CG M2; dark green dash dotted lines: CG M3; blue dotted lines: chl, backmapped from CG M2.

This journal is © the Owner Societies 2015

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

of the ring beads show a maximum which is too close to the polar head groups of DPPC compared to the AA reference. When a less polar bead type, SC3, is used for aromatic ring carbons in model M2, the distributions are shifted more towards the hydrophobic region of DPPC, thus corresponding better to the AA behavior. The partitioning is even better reproduced in model M3 since here we have used the less polar bead type, Na, for aldehyde/ketone/ester groups. This specific partitioning of the porphyrin ring within the bilayer is demonstrated in a CG chl *b* snapshot of the inset of Fig. 5(a). The AA chl *b* ring is drawn more deeply into the tail region than the CG ones, because of the fact that, atomistically, it can drag a few lipid molecules along. Similar observations have been found in other models for the distribution of amino acids in a lipid bilayer.[39] To explore the influence of the tail on the partitioning of chl within the bilayer, we calculate the density distributions of the porphyrin ring region of the model systems without tails, S3 and S4 (shown in the ESI,† in Fig. S4) using the parameters obtained from model M2. One observes a shift of the peak position of the density distribution of the chl ring without tails towards the polar head group region of the bilayer. Thus, we can conclude that the tail "pulls" the ring system of chl more into the apolar region of the bilayer.

Fig. 5(b) and (d) show the partitioning behavior of the tail of chl *b* and chl *a* within the DPPC bilayer, respectively. The first beads of the tail exhibit a polar ester group and an unsaturated double bond while the remaining region of the tail is hydrophobic. This causes the region where the tail is connected to the ring system to reside near the polar head group region of the DPPC bilayer. The remaining part of the tail favors to stay in the hydrophobic lipid tail region. The AA tail density is again better reproduced from model M3 than that from M2 and M1. The backmapped density distributions are broad in comparison to the CG and the AA one. The inset in Fig. 5(c) again depicts the location of the lipids and the chl *b* tail – showing how the chl tail resides in the hydrophobic lipid region.

4.3 Orientation of chl

The orientation of chl in the lipid bilayer is analyzed as a further indicator, whether the interaction of various CG beads of chl with the lipid components is reasonably represented and balanced in the CG model. To do so, we define a vector across the porphyrin ring that helps to quantify the orientation of chl with respect to the bilayer normal. As shown in Fig. 6 vector **v1** is defined between the ring beads SC3_6 and SC3_8, with SC3_6 being the bead closest to the tail and SC3_8 lying on the opposite side of the ring (corresponding to the CMC bead shown in Fig. 1).

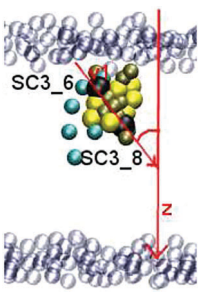

Fig. 6 Angle between vector **v1** within the porphyrin ring of chl (defined by the beads SC3_6 and SC3_8) and the bilayer normal that is used to analyze the orientation of chl in the bilayer.

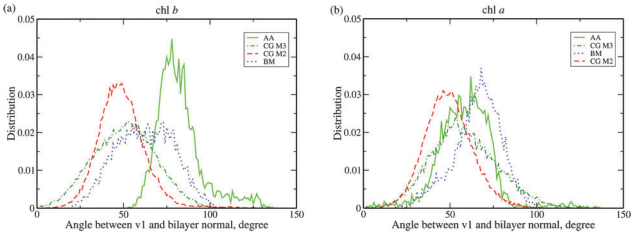

Fig. 7 Distributions of the angle between **v1** and bilayer normal (see Fig. 6) for (a) chl *b* and (b) chl *a*. Green solid lines: AA; red dashed lines: CG M2; dark green dash dotted lines: CG M3; blue dotted lines: AA backmapped from CG M2.

Fig. 6 illustrates a characteristic configuration including the angle between **v1** and the bilayer normal. Fig. 7(a) and (b) show the tilt angle distributions between **v1** and the bilayer normal for chl *b* and chl *a*, respectively. One should note that the angular distributions correspond to a single chl *b* or chl *a* molecule in the AA simulation while the CG distributions are averaged over 8 molecules. The width of the distributions is comparable for both the AA and the CG model, indirectly leading to the conclusion that the CG model is as flexible as the AA one. The agreement between CG and AA simulations is better for chl *a* than for chl *b*. This is due to the fact that the angular behavior of chl *b* is influenced by interactions of the two polar groups on opposite sides of the ring system with the lipid head groups (as discussed previously in Fig. 1). These rather specific interactions are probably not entirely captured in the CG model. Importantly, when the polarity of ester/aldehyde/ketone functional groups in the CG level is reduced from M2 to M3 by replacing bead type P2/3/4 with a less polar bead type, Na, the angular distributions from M3 shift closer to the AA one than that from M2 leading to the conclusion that M3 is better representing the interactions with the bilayer. After backmapping, however, the angular distributions revert back towards the ones from the original AA simulations.

As discussed previously, we have found that the tail plays a significant role in the partitioning and orientation of the rings within the bilayer. To understand the influence of the tail on the orientation of chl, we calculate the tilt angle of the model system without a tail and the real system with a tail. In general, the tilt angle for the model CG chl *b* system exhibits larger fluctuations than the real ones (see Fig. S5 in the ESI†) on longer timescales. Moreover, the tilt angle of the model CG chl *b* ring exhibits larger fluctuations than the model CG chl *a* ring. Presumably this is due to the presence of the polar groups on two opposite sides of the chl *b* ring, which interact with the polar head group region of the bilayer. This leads to an additional driving force to orient the ring perpendicular to the bilayer normal which is counteracting the natural tendency of the ring system to be rather oriented parallel to the bilayer normal (bringing the aromatic ring system in contact with the hydrophobic lipid tails and disturbing the order of the lipids less). The latter tendency can be observed for the chl *a* ring

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

system without tails – which does not have the additional polar aldehyde group.

4.4 Dynamics

Next, we calculate the translational diffusion in the *xy*-plane (MSD_{xy}), *i.e.*, in the plane perpendicular to the bilayer normal. Fig. 8(a) and (b) describe the MSD_{xy} for chl *b* and chl *a*, respectively. The diffusive regime is observed at times of 100 ns for the molecules with tails. The in-plane mean square displacements from our CG models, M2 and M3, follow similar behavior and the in-plane CG diffusion constant ranges approximately on the order of 10^{-9} cm^2 s^{-1} which is two orders of magnitude slower than that of the DPPC.[40] From the shift in the log–log plot of the MSD_{xy} *vs.* time one sees that the CG model is 7–8 times faster than the AA model. As the degrees of freedom are reduced from the AA to the CG model, the underlying CG energy surface is smoother than that of the AA one.[41,42] This reduces the frictional coefficient in the CG model[15] resulting in faster dynamics in the CG level compared to the AA one. The translational diffusion of chl in the presence and the absence of the tail does not differ much which indicates that the translational diffusion is dominated by the ring. As the preferred orientation of the ring is not exactly parallel to the lipid tails (as seen from the angular distributions in Fig. 7), the translational motion of the ring is highly hindered by the restricted environment within lipids. This inhibits fast translational motion of the molecules even if there is no tail present in the model system. Next, we calculate the rotational correlation function of the model chl ring and the full molecule to investigate the effect of the tail on the re-orientational dynamics. We take the cross-product of the two vectors across the ring connecting SC3_8, SC3_6, and SC3_6, SC3_4, generating a vector perpendicular to the plane of the ring of chl. We calculate the rotational auto-correlation function of this vector using a first order Legendre polynomial. Fig. 8(c) and (d) show the rotational autocorrelation function of chl *b* and chl *a* molecules, respectively depicting that the rotational dynamics of the model ring system without the tail decays faster than the real system with the tail – which differs from translational motion. This fits to the previous observation of restricted angular motion invoked by the tail (Fig. S5 in the ESI†).

4.5 Aggregation of chl in the lipid bilayer

The results in the previous subsections clearly show that the dynamics of the chl molecule in the lipid bilayer is very slow compared to the timescales of AA simulations, in particular in the case of the full molecule that is anchored in the bilayer by its tail. We see that the combination of CG simulations with AA ones after backmapping can to some extent alleviate this problem. Nevertheless, a full validation of the equivalence of the CG and AA simulation level is incredibly computationally extensive if not unfeasible. In particular we see that the timescale in which one can converge association of pigments coupled to reorientation is extremely long and would require the application of methods such as umbrella sampling with multiple umbrellas. Instead we have qualitatively analyzed the propensity of the chl pigments to aggregate in the lipid bilayer

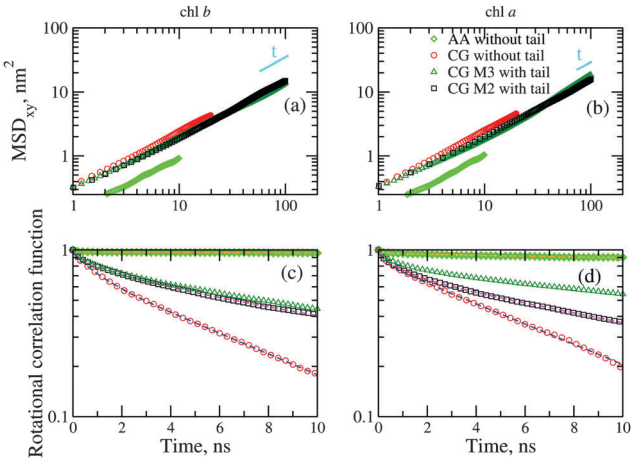

Fig. 8 Mean square displacements (MSD_{xy}) in the *xy* plane within the lipid bilayer (upper panels) and rotational autocorrelation functions (RACFs) (lower panels) of chl *b* (left panels) and chl *a* (right panels) as a function of simulation time. RACFs have been computed using a vector perpendicular to the porphyrin ring (see the text). Green symbols: AA molecules without the tail; red symbols: CG molecules (M2) without the tail; black symbols: CG molecules (M2) with the tail; dark green symbols: CG M3 with the tail; fine lines: fit of the RACF to extract relaxation times (see the text).

as a last aspect of validation of the CG model before we assess the applicability of the model to the protein/pigment complex.

We have performed an analysis of the formation of chl clusters (with a cutoff of 0.6 nm, which corresponds to the first peak of the Mg–Mg RDF and therefore gives a good estimate of a typical distance of two chl molecules in contact), taking care that we separately analyze clustering within the individual leaflets of the bilayer. Fig. 9(a) and (b) show the resulting distribution of cluster sizes (*i.e.* the number of individual chl molecules and clusters of 2, 3 and 4 molecules) obtained after averaging over both leaflets from three CG models (M1, M2 and M3). We have observed in all models a tendency to aggregate,

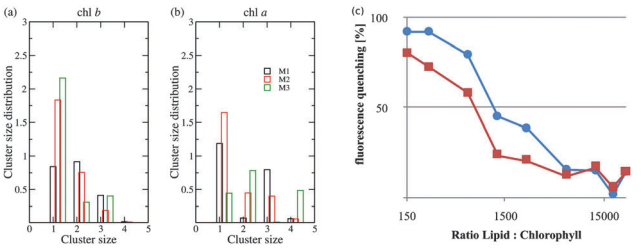

Fig. 9 Aggregate analysis of (a) chl *b* and (b) chl *a* for three models (M1, M2 and M3) averaged over two leaflets. (c) Fluorescence quenching of chl to monitor aggregation of chl in DPPG liposomes. Liposomes of about 100 nm diameter were made of 1,2-dipalmitoyl-*sn-glycero*-3-phospho-(1′-*rac*-glycerol) (DPPG) and chl (chl *a* : *b* = 1 : 1) at the given lipid-chlorophyll ratios (LCRs). The quenching of chl *b* fluorescence at 660 nm was measured in dependence of the LCR, relative to its fluorescence at a LCR of 25 000 where neither aggregation nor energy transfer from chl *b* to chl *a* is thought to take place. The total quenching of chl *b* fluorescence (blue circles) was corrected for the contribution to chl *b* fluorescence quenching of the energy transfer from chl *b* to chl *a* to monitor the amount of quenching due to chl aggregation (red squares). For a more detailed discussion and experimental details see ESI.†

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

and we have also found that the clusters form and break multiple times in the course of the simulation, *i.e.* the aggregation is not overly strong. Qualitatively, these data are corroborated by experimental observations. Upon aggregate formation, chlorophylls become non-fluorescent and dissipate their excitation energy as heat. This effect is thought to contribute to the protection of photosynthetic organisms from photodamage due to excess light in a process called non-photochemical quenching (NPQ).[43] Chl dissolved in organic solvents or in lipid membranes forms such non-fluorescent aggregates in a concentration-dependent manner.[44,45] The degree of fluorescence quenching gives an estimate of the percentage of chl molecules sequestered into aggregates. With a mixture of chl *a* and chl *b* dissolved in lipid vesicles of dipalmitoyl phosphatidylglycerol (DPPG), significant fluorescence quenching was observed if the lipid : chl ratio dropped below 1250 (Fig. 9(c)), which is consistent with other observations reported in the literature.[44] This indicates that at the lipid : chl ratio used in the model, chl aggregation is in fact to be expected.

4.6 Comparison of AA and CG pigmented LHCII trimers

We have simulated an AA LHCII in the presence of 512 DPPC and 26249 SPC/E water molecules. The initial configuration of the protein trimer is obtained from X-ray data (1RWT.pdb)[4] which is embedded in an already equilibrated DPPC bilayer using a g_membed[46] tool of gromacs. 33 sodium ions are added to make the system neutral and the system is energy minimized followed by a 200 ns *NPT* run with the same parameters as mentioned before. The final configuration of the AA LHCII in the DPPC bilayer is mapped into the CG representation. The CG force-field for the proteins involves an elastic network,[47] but there is no elastic network between the proteins and the chl molecules. The CG LHCII is simulated for 100 ns in an *NPT* ensemble preceded by energy minimization and equilibration with the same set of parameters as mentioned in Section 3.4. Unlike our initial attempts without the careful parameterization of the pigments, the trimeric CG protein–pigment complex has been stable now. The properties of the complex from the CG model are in good agreement with the AA reference simulation. In Fig. 10(c) and (d), a comparison of contact maps for the pigment-protein system in the CG simulation with the AA reference is shown. It can be seen that the pigments are in stable contact at the respective binding interfaces – most notably without the presence of any artificial elastic network between the protein core and the pigments. A more detailed study is in preparation.

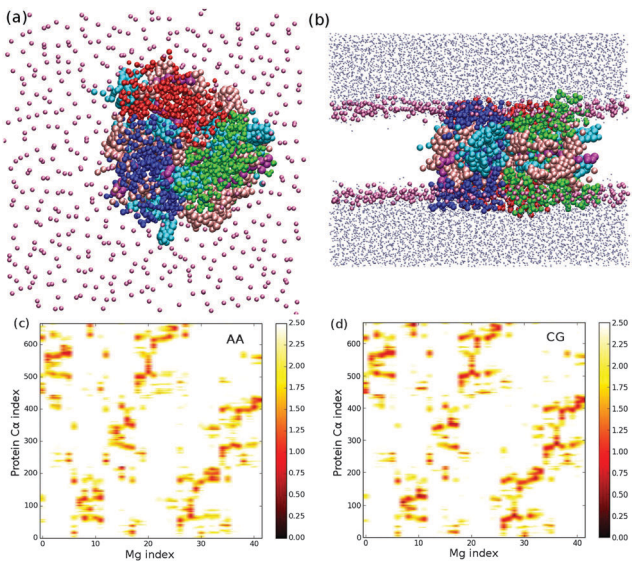

Fig. 10 Snapshots from a CG simulation of an LHCII trimer: (a) top and (b) side view in a simulation at CG resolution. Colors according to the chain or molecule type: blue – chain A; red – chain B; green – chain C; cyan – chl *b*, light pink – chl *a*, mauve – DPPC head groups, and light blue – water. Lower panels: contact maps between chl pigments and protein residues of the LHCII trimer – drawn as distance maps between the C_α atoms of the proteins (*y*-axis) and the Mg atoms of all chl pigments (*x*-axis) within a 2.5 nm cut-off for 70 ns long AA (left panel) and 100 ns long CG (right panel) simulations using model M3. The maps show that the pigments are stably located in their binding sites for both levels of resolution.

5 Conclusions and outlook

We have derived a CG model for chl *b* and *a* in the DPPC bilayer based on the combination of a structure-based approach for bonded and a mixed structure-based and partitioning-based approach for non-bonded interaction potentials to fit the MARTINI force-field. Boltzmann inversion of the atomistic bonded distributions has generated a final CG set of parameters that reproduces the behavior of the AA system well. Also distances and angles between CG beads, which are not immediately connected and have, therefore, not been part of the parameterization procedure, are reproduced very well. Thus, the overall shape of the porphyrin ring and the different conformations of the phytol tail are well represented in this CG model. The non-bonded potentials of the presented CG model produce the correct partitioning of chl within the bilayer which is found to be strongly affected by the chemical interactions between chl and the lipids. We found that the porphyrin ring is preferentially located in the hydrophobic region of the bilayer, but not in the center, rather close towards the polar interfaces dictated by balance between polar–polar interactions and hydrophobic interactions between chl and lipid molecules. The orientation of the chl within the bilayer also depends on the chemical nature of the molecule and their interactions with the lipids. The tilt angle distributions from the AA reference are well reproduced by the CG model. Note, that the different physical properties of chl, such as partitioning within the bilayer or the orientation of chl, do not only depend on the parameters of chl, but also strongly on the lipid parameters. As we use MARTINI parameters for the lipid molecules, possible limitations of the lipid force-field affect the behavior of CG chl. An adaptation of the lipid parameters, however, appears to be not useful at the present state since that would interfere with other interactions (for example with protein components) within the well-established and well-tested force-field. The main target of the present paper was to derive parameters for the pigments that are well-compatible with the MARTINI lipid and protein forcefields with the well-tested

Open Access Article. Published on 22 July 2015. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

DPPC lipid reference system. We expect that these parameters will be transferable to other lipids in the MARTINI forcefield.

Translational and rotational diffusion of chl (with and without the phytol tail) in the bilayer have been investigated. Even though the translational diffusion is 7–8 times faster in the CG simulation compared to the AA one, it is in general very slow in nature irrespective of the presence of the phytol tail. This is different for rotational reorientation of the ring system which is much more restricted by the phytol tail than the translational motion – probably because the tail anchors and orients the chl molecule within the lipids.

From a long CG simulation AA details are reintroduced by backmapping and the structural properties of the original AA simulation are recovered. This indicates that those types of structural properties, which require chemical details but are not accessible in AA simulations alone due to time-scale limitations, can be obtained with this switching back and forth between different levels of resolutions. This will be applied to study the light harvesting protein–pigment complex in the future. Unlike our initial attempts without the careful parameterization of the pigments, a first simulation of the LHCII complex with the final CG chl model has shown to be very promising. The properties of the complex from the CG model are in good agreement with the atomistic ones. Using this CG model, we aim to study the nature of various properties of the light harvesting complex in the future.

Acknowledgements

This collaboration has been carried out within the Collaborative Research Center (SFB 625) in Mainz: "From Single Molecules to Nanoscopically Structured Materials". We would like to thank Shahua Ding, Sandra Ritz, and Eva Sinner for many inspiring discussions. We would also like to thank Mara Jochum and Biswaroop Mukherjee for careful reading of this manuscript and Christoph Globisch, Denis Andrienko, Victor Ruehle and Christoph Junghans for their technical help. AD acknowledges financial support from the fellowship program of the Max Planck Society, and CP is very grateful for support from the German Research Foundation within the Emmy Noether program.

References

1 B. van Oorta, A. van Hoeka, A. V. Rubanb and H. van Amerongen, *FEBS Lett.*, 2007, **581**, 3528.
2 C. Dockter, A. Volkov, C. Bauer, Y. Polyhach, Z. Joly-Lopez, G. Jeschke and H. Paulsen, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 18485.
3 J. P. Dekker, H. van Roon and E. J. Boekem, *FEBS Lett.*, 1999, **449**, 211.
4 Z. Liu, H. Yan, K. Wang, T. Kuang, J. Zhang, L. Gui, X. An and W. Chang, *Nature*, 2004, **428**, 287.
5 J. Standfuss, A. C. Terwisscha van Scheltinga, M. Lamborghini and W. Kühlbrandt, *Org. Biomol. Chem.*, 2005, **24**, 919.
6 J. Linnanto and J. Korppi-Tommola, *Phys. Chem. Chem. Phys.*, 2006, **8**, 663.
7 O. Carsten and K. Ulrich, *J. Phys. Chem. B*, 2010, **114**, 12427.
8 K. Karki and D. Roccatano, *J. Chem. Theory Comput.*, 2011, **7**, 1131.
9 G. S. Ayton, W. G. Noid and G. A. Voth, *Curr. Opin. Struct. Biol.*, 2007, **17**, 192.
10 S. Tanaka and H. A. Scheraga, *Macromolecules*, 1976, **9**, 945.
11 S. Miyazawa and R. L. Jernigan, *Macromolecules*, 1985, **18**, 534.
12 A. Liwo, A. Czaplewski, J. Pillardy and H. A. Scheraga, *J. Chem. Phys.*, 2001, **115**, 2323.
13 T. Bereau, Z.-J. Wang and M. Deserno, *J. Chem. Phys.*, 2014, **140**, 115101.
14 S. J. Marrink, A. H. de Vries and A. E. Mark, *J. Phys. Chem. B*, 2004, **108**, 750.
15 S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Phys. Chem. B*, 2007, **111**, 7812.
16 L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman and S.-J. Marrink, *J. Chem. Theory Comput.*, 2008, **4**, 819.
17 F. J. van Eerden, D. H. de Jong, A. H. de Vries, T. A. Wassenaar and S. J. Marrink, *Biochim. Biophys. Acta, Biomembr.*, 2015, **1848**, 1319.
18 E. Villa, A. Balaeff and K. Schulten, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6783.
19 E. Lyman, F. M. Ytreberg and D. M. Zuckerman, *Phys. Rev. Lett.*, 2006, **96**, 028105.
20 R. Potestio, P. Español, R. Delgado-Buscalioni, R. Everaers, K. Kremer and D. Donadio, *Phys. Rev. Lett.*, 2013, **111**, 060601.
21 M. Orsi, W. E. Sanderson and J. W. Essex, *J. Phys. Chem. B*, 2009, **113**, 12019.
22 T. A. Wassenaar, H. I. Ingòlfsson, M. Prie, S. J. Marrink and L. V. Schäfer, *J. Phys. Chem. B*, 2013, **117**, 3516.
23 A. R. Leach, *Molecular modelling: principles and applications*, Pearson Education Limited, Essex, England, 2001.
24 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
25 M. Parinello and A. Rahman, *J. Appl. Physiol.*, 1981, **52**, 7182.
26 H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43.
27 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model.*, 2001, **7**, 306.
28 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435.
29 A. Kukol, *J. Chem. Theory Comput.*, 2009, **5**, 615.
30 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269.
31 D. Fritz, V. A. Harmandaris, K. Kremer and N. F. A. van der Vegt, *Macromolecules*, 2009, **42**, 7579.
32 N. F. A. van der Vegt, C. Peter and K. Kremer, *Coarse-graining of Condensed Phase and Biomolecular Systems*, Chapman and Hall/CRC Press, 2008.
33 O. Engin, A. Villa, C. Peter and M. Sayar, *Macromol. Theory Simul.*, 2011, **20**, 451.

34 O. Bezkorovaynaya, A. Lukyanov, K. Kremer and C. Peter, *J. Comput. Chem.*, 2012, **33**, 937.
35 B. Mukherjee, L. Delle Site, K. Kremer and C. Peter, *J. Phys. Chem. B*, 2012, **116**, 8474.
36 V. Rühle, C. Junghans, A. Lukyanov, K. Kremer and D. Andrienko, *J. Chem. Theory Comput.*, 2009, **5**, 3211.
37 A. J. Rzepiela, L. V. Schäfer, N. Goga, H. J. Risselada, A. H. De Vries and S. J. Marrink, *J. Comput. Chem.*, 2010, **31**, 1333.
38 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701.
39 J. L. MacCallum, W. F. D. Bennett and D. P. Tieleman, *Biophys. J.*, 2008, **94**, 3393.
40 M. Patra, M. Karttunen, M. T. Hyvnen, E. Falck and I. Vattulainen, *J. Phys. Chem. B*, 2004, **108**, 4485.
41 V. A. Harmandaris and K. Kremer, *Macromolecules*, 2009, **42**, 791.
42 D. Fritz, K. Koschke, V. A. Harmandaris, N. F. A. van der Vegt and K. Kremer, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10412.
43 R. Goss and B. Lepetit, *J. Plant Physiol.*, 2015, **172**, 13.
44 A. G. Lee, *Biochemistry*, 1975, **1787**, 4397.
45 G. Beddard, S. Carlin and G. Porter, *Chem. Phys. Lett.*, 1976, **43**, 27.
46 M. W. M. Hoefling, C. Aponte-Santamaría, H. Grubmüller and G. Groenhof, *J. Comput. Chem.*, 2010, **31**, 2169.
47 X. Periole, M. Cavalli, S.-J. Marrink and M. A. Ceruso, *J. Chem. Theory Comput.*, 2009, **5**, 2531.
48 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33.

Open Access Article. Published on 22 July 2015.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.