

# Energy Flow in the Cryptophyte PE545 Antenna is Directed by Bilin Pigment Conformation

*Carles Curutchet,<sup>\*,a</sup> Vladimir I. Novoderezhkin,<sup>b</sup> Jacob Kongsted,<sup>c</sup>  
Aurora Muñoz-Losa,<sup>de</sup> Rienk van Grondelle,<sup>f</sup> Gregory D. Scholes,<sup>g</sup>  
and Benedetta Mennucci<sup>d</sup>*

<sup>a</sup>Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Av. Joan XXIII s/n, 08028 Barcelona, Spain

<sup>b</sup>A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, Russia

<sup>c</sup>Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

<sup>d</sup>Dipartimento di Chimica e Chimica Industriale, Università di Pisa, via Risorgimento 35, 56126 Pisa, Italy

<sup>e</sup>Present adress: Área de Química Física, Universidad de Extremadura, Avda de Elvas s/n, 06071 Badajoz, Spain

<sup>f</sup>Department of Physics and Astronomy, Faculty of Sciences, VU University Amsterdam, Amsterdam, The Netherlands

<sup>g</sup>Department of Chemistry, 80 St. George Street, Institute for Optical Sciences, and Centre for Quantum Information and Quantum Control, University of Toronto, Toronto, Ontario, M5S 3H6 Canada

<sup>\*</sup>To whom correspondence may be addressed. E-mail: carles.curutchet@ub.edu

## ABSTRACT

Structure-based calculations are combined with quantitative modelling of spectra and energy transfer dynamics to determine the energy transfer scheme of the PE545 principal light-harvesting antenna of the cryptomonad *Rhodomonas* CS24. We use a recently developed quantum-mechanics/molecular mechanics (QM/MM) method that allows us to account for pigment-protein interactions at atomic detail in site energies, transition dipole moments and electronic couplings. In addition, conformational flexibility of the pigment-protein complex is accounted for through molecular dynamics (MD) simulations. We find that conformational disorder largely smoothes the large energetic

differences predicted from the crystal structure between the pseudosymmetric pairs  $\text{PEB}_{50/61\text{C}} - \text{PEB}_{50/61\text{D}}$  and  $\text{PEB}_{82\text{C}} - \text{PEB}_{82\text{D}}$ . Moreover, we find that in contrast to chlorophyll-based photosynthetic complexes, pigment composition and conformation play the major role in defining the energy ladder in the PE545 complex, rather than specific pigment-protein interactions. This is explained by the remarkable conformational flexibility of the 8 bilin pigments in PE545, characterized by a quasi-linear arrangement of 4 pyrrole units. The MD-QM/MM site energies allow us to reproduce the main features of the spectra, and minor adjustments of the energies of the three red-most pigments  $\text{DBV}_{19\text{A}}$ ,  $\text{DBV}_{19\text{B}}$ , and  $\text{PEB}_{82\text{D}}$  allow us to model the spectra of PE545 with a similar quality compared to our original model (Model E from Novoderezhkin et al., *Biophys. J.* 2010, 99, 344), which was extracted from the spectral and kinetic fit. Moreover, the fit of the transient absorption kinetics is even better in the new structure-based model. The largest difference between our previous and present results is that the MD-QM/MM calculations predict a much smaller gap between the  $\text{PEB}_{50/61\text{C}}$  and  $\text{PEB}_{50/61\text{D}}$  sites, in better accord with chemical intuition. We conclude that the current adjusted MD-QM/MM energies are more reliable in order to explore the spectral properties and energy transfer dynamics in the PE545 complex.

**KEYWORDS:** Pigment–protein complex, Electronic energy transfer, Light harvesting, Modified Redfield Theory, Site energies.

## 1. Introduction

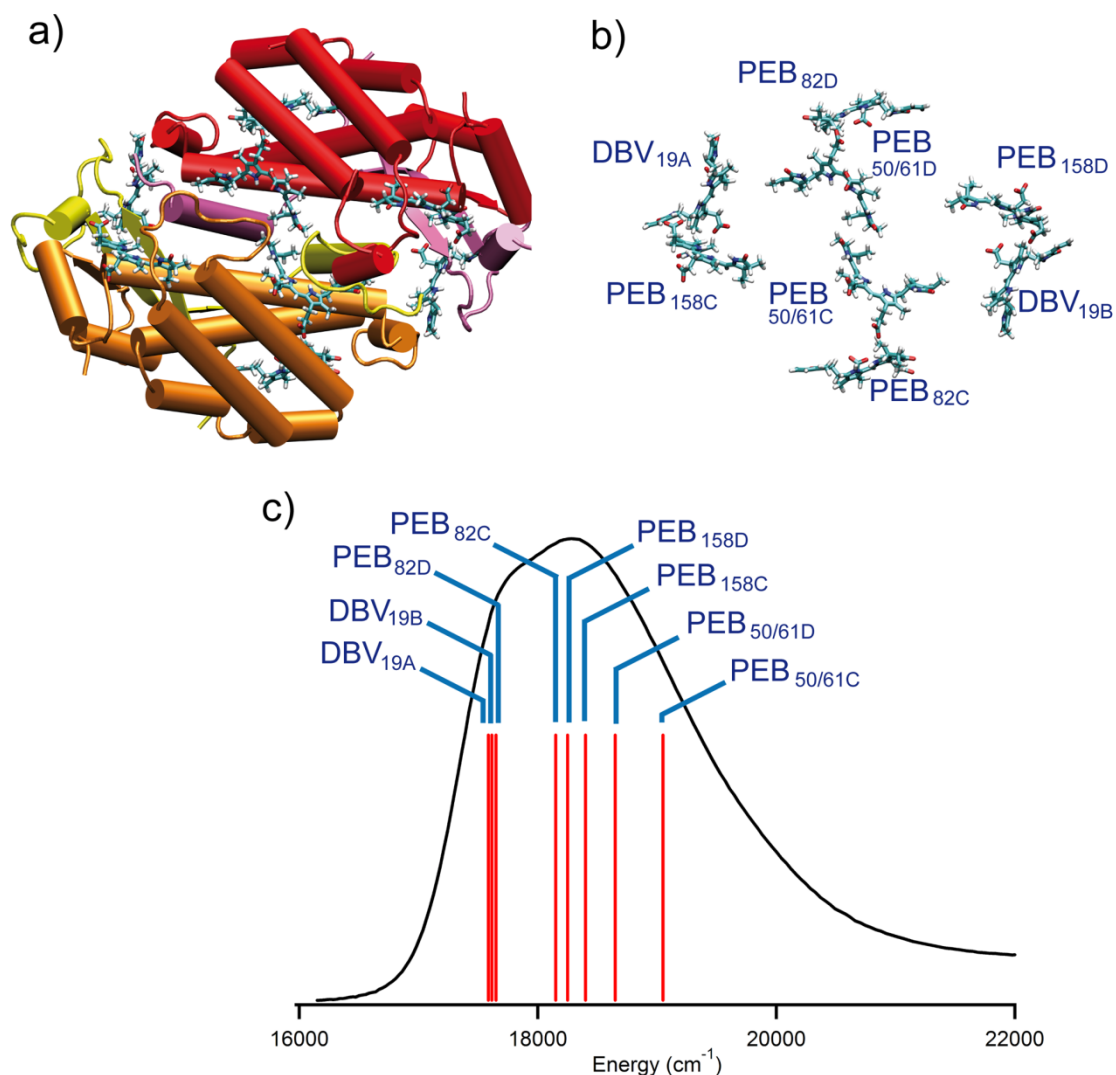
Photosynthetic antennae are special proteins characterized by carrying a large number of light-absorbing molecules (pigments), aimed at capturing sunlight and transferring this energy through a number of electronic energy transfer (EET) steps toward reaction centers, where the energy is subsequently stored through a number of chemical reactions.<sup>1-2</sup> This principle inspires the quest toward practical artificial photosynthetic devices, in which sunlight could be directly stored in the form of chemical fuels, thus lowering our dependency on fossil fuels.<sup>3-4</sup> Because light-harvesting is performed with very high quantum efficiencies (>90%) in natural biological systems, a detailed understanding of the mechanisms of photosynthetic EET has been of interest for a long time. A key objective in this area of research consists of developing integrated approaches able to describe the relation between structure and function of the light-harvesting antenna.<sup>5-6</sup>

The fundamental properties of such systems can, to some degree, be understood based on Förster theory, which predicts incoherent EET dynamics according to the dipole-dipole interactions between pairs of pigments in the complex, as well as the spectral overlap between their optical line shapes.<sup>7</sup> However, a microscopic understanding of the role played by the environment is necessary in order to have a complete picture of how structure dictates EET dynamics.<sup>8</sup> In particular, for non-Förster energy transfer mechanisms it has become apparent that much more detailed insights into how molecules interact with complex environments are needed.<sup>9-10</sup> For instance, the protein environment actively modulates the dipole-dipole interactions, an effect that is described in Förster theory through a crude screening factor  $1/n^2$ , where  $n$  is the refractive index of the system. Recent developments of quantum chemical methods coupled to continuum or quantum mechanical/molecular mechanical (QM/MM) descriptions of the environment have shown that screening effects sensitively depend on the distance and mutual arrangement between the molecules,<sup>11-12</sup> as well as on the local protein dielectric environment that surrounds them.<sup>13</sup>

Another key role played by the environment consists of the time-dependent modulation of the so-called site energies, i.e. the uncoupled electronic transition levels localized on each pigment. Such pigment-protein coupling, for instance, dictates line

shapes, decoherence and energy transfer, and can be described through the frequency spectrum of the environment (spectral density). Equally important, though, is the magnitude of the pigment-protein coupling, which individually shifts each site energy in the complex thus defining the energy pathways in the system. A common strategy to unravel the site energies of a pigment-protein complex consists of a simultaneous fitting of several optical spectra, while adopting a set of electronic couplings calculated from the X-ray structure.<sup>2, 14</sup> Such a strategy, however, becomes impractical when the number of pigments is very large, for instance, when dealing with the 96 chlorophyll a pigments in photosystem I (PSI).<sup>15</sup> A natural alternative consists in the calculation of site energies using quantum chemical approaches. In that case, though, errors can be substantial, so quantitative modeling of spectra and kinetics should be performed in order to elucidate a consensus model compatible with simulations and experiment.

Recently, Renger and co-workers have successfully applied a methodology that combines quantum chemistry and electrostatic calculations to unravel the site energies in a variety of chlorophyll-containing pigment-protein complexes.<sup>15-17</sup> Such a methodology derives energy shifts from the differential interaction of ground and excited state charge densities with the protein charge distribution, which is described by point charges, whereas protein polarization is tackled through a continuum dielectric model. This approach has proven to be quite robust. However, it neglects how the protein modulates the chromophore conformation, an effect that further tunes its electronic transition energy. Although this effect can be minor when dealing with chlorophylls, pigments characterized by a semi-rigid central porphyrin ring, it is expected to play a more significant role in bilins, characterized by a quasi-linear arrangement of four pyrrole rings.<sup>18</sup> Recently, other groups have computed the site energies in the Fenna-Matthews-Olson (FMO) complex by combining quantum chemical calculations with molecular dynamics (MD) simulations, thus allowing for conformational flexibility in the complex.<sup>19-20</sup> The results do not differ much from the previous set derived by Renger and co-workers,<sup>16</sup> thus apparently confirming that protein-induced distortions of the pigment conformation are not essential to describe chlorophyll site-energies.



**Figure 1.** (a) Structure of the PE545 light-harvesting complex. (b) Detailed view of the eight light-absorbing bilin molecules. (c) Electronic absorption spectrum of the isolated PE545 protein in aqueous buffer (294 K) with approximate absorption peaks corresponding to the bilin molecules.

In this work we address the determination of site energies in a rather challenging system, the PE545 principal light-harvesting antenna of the cryptomonad *Rhodomonas* sp. strain CS24.<sup>13, 18, 21-22</sup> The crystal structure of PE545, illustrated in Figure 1, is organized as an  $\alpha_1\alpha_2\beta\beta$  dimer, and has been determined at 1.63 Å<sup>23</sup> and later at 0.97 Å resolution<sup>21</sup>. It contains 8 bilin chromophores, characterized by a linear tetrapyrrole structure covalently linked to the protein scaffold. In particular, each  $\alpha$  chain (A and B) contains a 15,16-dihydrobiliverdin (DBV), whereas each  $\beta$  polypeptide chain (C or D) is linked to three phycoerythrobilins (PEB). The corresponding chromophores are

labeled  $\text{DBV}_{19\text{A}}$ ,  $\text{DBV}_{19\text{B}}$ ,  $\text{PEB}_{158\text{C}}$ ,  $\text{PEB}_{158\text{D}}$ ,  $\text{PEB}_{50/61\text{C}}$ ,  $\text{PEB}_{50/61\text{D}}$ ,  $\text{PEB}_{82\text{C}}$  and  $\text{PEB}_{82\text{D}}$ , where the subscript denotes the protein subunit and cysteine residue linked to the chromophore. The central  $\text{PEB}_{50/61}$  pigments are linked to the protein by two cysteine residues. The overall PE545 structure (and the chromophores) displays a pseudosymmetry about the 2-fold axis relating the  $\alpha_1\beta$  and  $\alpha_2\beta$  monomers.

As mentioned above, bilins are linear tetrapyrroles that have a larger degree of conformational flexibility compared to chlorophylls. Thus, modulation of the site energies by the protein scaffold could be achieved by constraining the conformational space of the pigments. In addition, the DBVs have an extended conjugated  $\pi$ -system that red-shifts their spectral properties compared to PEBs, thus making them the obvious candidates for the energy trapping site in the complex. Previous theoretical studies suggested an energy ordering  $\text{DBV}_{19} < \text{PEB}_{82} < \text{PEB}_{158} < \text{PEB}_{50/61}$  in the complex<sup>12, 18</sup>. More recently, we modeled steady-state spectra and transient absorption of PE545 with modified Redfield theory.<sup>22</sup> The results suggested that quite significant energetic differences between pseudosymmetric pigments are needed in order to describe the spectroscopic data for the complex. However, several sets of site energies were found to be compatible with the spectra and kinetics. This translates into several unanswered questions, for instance, about the relative energetic ordering of the  $\text{PEB}_{82\text{C}}$ ,  $\text{DBV}_{19\text{A}}$  and  $\text{DBV}_{19\text{B}}$  lowest energy chromophores.

In the present study we combine the MD-QM/MM methodology we recently developed<sup>24</sup> with spectral and kinetics modeling in order to determine a consensus model of the PE545 light-harvesting pathways. This methodology allows us to consistently incorporate environmental effects (electrostatic and polarization) at atomic detail not only in the evaluation of site energies, but also in the description of screening effects in the electronic couplings. From the calculated parameters, we model the spectra and excitation dynamics in PE545, and find a reasonable agreement with experiment, which can be further improved by minor adjustments of the calculated site energies. Our proposed model differs from the previous set of energies derived by fitting to experiments.<sup>22</sup> We find that the energetic differences between pseudosymmetric pairs, like  $\text{PEB}_{50/61\text{C}}/\text{PEB}_{50/61\text{D}}$  or  $\text{PEB}_{82\text{C}}/\text{PEB}_{82\text{D}}$ , are significantly smaller than previously suggested, in better accord with chemical intuition. In addition, we find that energy tuning in PE545 is mostly achieved by constraining the

conformational space available to the bilin pigments in the protein scaffold, rather than by specific pigment-protein interactions.

This paper is organized as follows. In Section 2 we describe the methods and computational details of the structure-based calculations as well as the modeling of spectra and kinetics. In Section 3.1, first we analyze the numerical robustness of our QM/MMpol methodology toward a number of issues. Then, in Section 3.2 we analyze the impact of conformational flexibility, as well as specific pigment-protein interactions, in tuning the energy levels of PE545. In Section 3.3 we model the spectra and kinetics of PE545 based on our calculated parameters and critically compare the results to experiments, with the aim of determining a consensus set of values for the PE545 site energies. Finally, we provide the conclusions and some future perspectives.

## 2. Methods

### 2.1 QM/MMpol model for energy transfer

In conventional QM/MM approaches, the environment is treated by assigning partial point charges to the atomic sites and the potential due to these point charges is introduced into the Hamiltonian of the QM system to give an effective Hamiltonian:

$$\hat{H}_{eff}|\Psi\rangle = (\hat{H}_0 + \hat{H}_{env})|\Psi\rangle = E|\Psi\rangle \quad (1)$$

Where  $\hat{H}_0$  is the Hamiltonian of the isolated QM system (here the chromophores), and the operator  $\hat{H}_{env}$  introduces the coupling between the chromophores and the environment. The addition of  $\hat{H}_{env}$  to the Hamiltonian automatically leads to a modification of the chromophore wavefunction, which has now to be determined by solving the effective equation (1). This can be done using the same methods developed for isolated molecules; here in particular we shall focus on the standard Self-Consistent-Field (SCF) approach.

In our QM/MM approach, this standard MM description is improved by introducing chromophore-environment mutual polarization effects in terms of atomic polarizabilities which give rise to an induced charge distribution described by induced dipoles; as a result  $\hat{H}_{env}$  can be split into four terms

$$\hat{H}_{env} = \hat{H}_{QM/MM} + \hat{H}_{MM} = \hat{H}_{QM/MM}^{ele} + \hat{H}_{QM/MM}^{pol} + \hat{H}_{MM}^{ele} + \hat{H}_{MM}^{pol} \quad (2)$$

The first term in Eq. (2),  $\hat{H}_{QM/MM}^{ele}$ , describes the electrostatic interaction between the QM system and the atomic point charges in the MM region, whereas the second term,  $\hat{H}_{QM/MM}^{pol}$ , represents the polarization interaction between the induced dipole moments and the electric field from the QM system. Finally, the MM term contains both the electrostatic self-energy of the MM charges ( $\hat{H}_{MM}^{ele}$ ), and the polarization interaction ( $\hat{H}_{MM}^{pol}$ ) between such charges and the induced dipoles. We note that the  $\hat{H}_{MM}^{ele}$  term enters in the effective Hamiltonian only as a constant energetic quantity, while the  $\hat{H}_{MM}^{pol}$  contribution depends on the QM wavefunction through the induced dipoles. In addition, we do not consider short-range dispersion and repulsion contributions in  $\hat{H}_{env}$ , since, as in most combined QM/MM methods, these are described by empirical potentials independent of the QM electronic degrees of freedom, thus not affecting our results.

The induced dipoles are dissected into a nuclear ( $n$ ) and an electronic ( $e$ ) component, namely

$$\vec{\mu}_i^n = \alpha_i \vec{E}_i(Z_n, q_m, \vec{\mu}_{j \neq i}^n) \quad (3)$$

$$\vec{\mu}_i^e = \alpha_i \vec{E}_i(\rho_e, \vec{\mu}_{j \neq i}^e)$$

where we have assumed a linear approximation, neglected any contribution of magnetic character related to the total electric field, and used an isotropic polarizability ( $\alpha_i$ ) for each selected point in the MM part of the system. The nuclear component is computed only once as it depends on the electric field originated by the nuclear charges ( $Z_n$ ), the MM charges ( $q_m$ ) and the nuclei-induced dipoles ( $\vec{\mu}_i^n$ ). On the contrary, the electronic component depends on the SCF density ( $\rho_e$ ) so it has to be updated in each SCF iteration. In Eq. (3), the electric field always contains contributions from the chromophore charge distribution, and the point charges and induced dipole moments in the MM region. This means that Eq. (3) must be solved iteratively, or, alternatively, it may be reformulated into a matrix equation,  $\mu^x = \mathbf{B}\mathbf{E}^x$ , where  $\mathbf{B}$  is a matrix of



dimension  $3N \otimes 3N$ , where  $N$  is the number of polarizable sites, which is determined uniquely by the positions of these sites and their polarizability values.

In this framework, the Fock operator which determines the SCF wavefunction of the solvated system becomes

$$\mathbf{F} = \mathbf{F}^0(\mathbf{P}) + \mathbf{h}^{ele} + \mathbf{h}^{pol} + \mathbf{X}^{pol}(\mathbf{P}) \quad (4)$$

where we have adopted the common expansion over a finite basis set, so all operators are given in a matrix form in such a basis, and  $\mathbf{P}$  represents the one-electron density matrix in the same basis set. The first term in Eq. (4) corresponds to the gas phase Fock matrix whereas the solvent-induced one- and two-electron operators are given by:

$$\begin{aligned} h_{\mu\nu}^{ele} &= \sum_i^M q_i V_{\mu\nu}(r_i) \\ h_{\mu\nu}^{pol} &= - \sum_i^N \vec{\mu}_i^n \cdot \vec{E}_{\mu\nu}(r_i) \end{aligned} \quad (5)$$

$$X_{\mu\nu}^{pol}(\mathbf{P}) = - \sum_i^N \vec{\mu}_i^e(\mathbf{P}) \cdot \vec{E}_{\mu\nu}(r_i)$$

where  $V_{\mu\nu}$  and  $E_{\mu\nu}$  indicate the electrostatic potential and electric field integrals.

This SCF-QM/MMPol approach can be further extended to a CIS formulation to calculate excited states and transition properties in a straightforward way once the proper Fock operator (4) has been defined. The CIS equations to be solved become

$$\begin{aligned} \mathbf{A}\mathbf{X}_n &= \omega_n \mathbf{X}_n \\ A_{ai,bj} &= \delta_{ab} \delta_{ij} (\epsilon_a - \epsilon_i) + \langle ia || bj \rangle + C_{ai,bj}^{pol} \end{aligned} \quad (6)$$

$$C_{ai,bj}^{pol} = - \sum_k^N \left( \int d\vec{r} \varphi_i(\vec{r}) \varphi_a^*(\vec{r}) \frac{(\vec{r}_k - \vec{r})}{|\vec{r}_k - \vec{r}|^3} \right) \cdot \vec{\mu}_k^e(\varphi_j \varphi_b^*)$$

where  $i,j$  and  $a,b$  indicate occupied and virtual orbitals, respectively.

We note that in the CIS matrix  $\mathbf{A}$ , the environment enters in two ways. First, the orbitals and the orbital energies have been modified by the environment through the addition of

the terms reported in Eq. (5) to the ground-state Fock operator and, secondly, a new environment-induced response term  $C$  is introduced to account for the change in the induced dipoles due to the excitation process. As a result, the eigenvectors  $\mathbf{X}_n$  and the eigenvalues  $\omega_n$  give now the transition density and transition energies for the molecule treated using QM in the presence of a polarizable MM environment. As reported in ref. [24], the transition densities obtained from solving Eq. (6) can be further used to compute the electronic coupling which defines the EET rate so that both implicit and explicit (or screening) effects due to the solvent are automatically taken into account: all the details of such a formulation can be found in the cited reference.

## 2.2 Computational details

In our approach, we followed a two-step strategy. First, we performed ground-state classical MD simulations adopting the parm99SB<sup>25-27</sup> non-polarizable force field (FF) from the Amber package. Compared to polarizable FFs, parm99SB has been extensively tested and balanced in order to correctly describe the structural dynamics of solvated proteins. As a second step, we performed QM/MM calculations of the EET parameters, but in this case switching to an explicit polarizable FF to describe the MM region. This latter choice arises from the need to differentiate between nuclear and electronic polarization effects when describing electronic transitions and electronic couplings.<sup>24</sup> We note here that the aim of the MD simulation is to sample the ground-state ensemble of the system in order to estimate statistically averaged EET parameters corresponding to the ground-state configuration. The details of the MD simulation of PE545 in water can be found in our previous publication.<sup>13</sup> From the trajectory, a total of 141 snapshots (extracted every 50 ps during the last 7 ns) were considered for QM/MM calculations.

In the MMPol calculations, the protein residues, solvent molecules, and bilin chromophores not included in the QM region, were described using atom centered charges and isotropic polarizabilities. The corresponding polarizable FF was derived from DFT calculations using the scheme outlined in Ref. [28]. In brief, this strategy consists in cutting the protein into single-residues capped with COCH<sub>3</sub> and NHCH<sub>3</sub> groups. Then, the residues are subjected to separate DFT calculations of the FF parameters. Atomic isotropic polarizabilities were calculated at the DFT(B3LYP)/aug-cc-pVDZ level using the LoProp<sup>29</sup> approach as implemented in the Molcas code,<sup>30</sup>

whereas atomic point charges were obtained from DFT(B3LYP)/cc-pVTZ ESP calculations followed by RESP constrained fittings as implemented in the Gaussian09<sup>31</sup> and Amber9<sup>32</sup> programs, respectively. Charges and polarizabilities were derived corresponding to the initial crystal structure and then used in all QM/MM calculations. The parameters for water, derived using the same strategy, were taken from Ref. [24].

In order to test the robustness of our approach we also used two alternative polarizable FFs, based on a completely different derivation of the parameters. First, we tested the Amber ff02 protein FF,<sup>33</sup> together with the water POL3 model,<sup>34</sup> which combines RESP charges with a modified version of Applequist polarization model. Applequist distributed polarizability values are fitted to reproduce a set of molecular experimental polarizabilities by adopting a full interactive model.<sup>35</sup> However, the Amber ff02 energy model, as well as the water POL3 model, neglects 1–2 and 1–3 interactions. This translates into a well-known systematic underestimation of polarization terms in ff02 and POL3, but unfortunately 1–2 and 1–3 interactions cannot be simply switched on, because this would lead to a problem known as the "polarization catastrophe", originated by an unphysical mutual overpolarization of closely-spaced dipoles.<sup>36-37</sup> Thole introduced an alternative strategy to avoid this problem by replacing the point dipoles by smeared charge distributions.<sup>38</sup> We also tested this strategy by adopting the atomic polarizabilities derived by van Duijnen and Swart,<sup>39</sup> which were obtained by fitting molecular polarizabilities to experiment using Thole's modified dipole-dipole tensor. In this case, we combined this polarization model with our RESP charge model. Note that a significant difference between these methods is the fact that LoProp polarizabilities are obtained for each individual MM site in the PE545 complex, whereas Applequist and van Duijnen values are assigned solely depending on the corresponding MM element, thus assuming transferability between different chemical environments.

Finally, in order to assess the role of the protein in tuning the site energies in the complex, additional calculations were performed using the Polarizable Continuum Model (PCM).<sup>40-41</sup> In the PCM calculations the environment was modeled as a dielectric continuum with a relative static dielectric constant of 15 and optical dielectric constant of 2.<sup>42</sup> The PCM cavities enclosing the chromophores were obtained in terms of interlocking spheres centered on selected nuclei. The chosen radii were obtained by

applying the United Atom Topological Model to the atomic radii of the UFF force field<sup>43</sup> as implemented in the Gaussian 09 code. All PCM and MMPol calculations were performed at the CIS/6-31G level of theory using a locally modified version of Gaussian09.<sup>31</sup>

## **2.3 Modeling of spectra and kinetics**

In our previous paper<sup>22</sup> we have modeled the spectra and kinetics of PE545 using modified Redfield theory. This exciton model included the static disorder (accounted for by introducing uncorrelated shifts of the site energies) and coupling to phonons (modeled by two overdamped Brownian oscillators and 14 underdamped high-frequency vibrations). The transition dipoles, center coordinates, and pigment-pigment couplings have been calculated using the PCM method, whereas the site energies (treated as free parameters) have been determined from a simultaneous fit of the absorption (OD), circular dichroism (CD), fluorescence (FL), excitation anisotropy (EA) spectra, and transient absorption (TA) kinetics. In the present study we calculate the spectra and kinetics for PE545 using the same approach, but now the site energies are taken from the MD-averaged MMPol calculations. The transition dipoles, and couplings are taken from the MMPol calculations as well. Expressions needed to calculate the OD, CD, FL, EA spectra and TA kinetics can be found in the Supporting Material of Ref. [<sup>22</sup>].

## **3. Results and discussion**

### **3.1 Validation and optimization of the QM/MMPol method**

The QM/MM scheme we use allows a rich description of the effect of the protein and solvent environment on the PE545 light-harvesting properties. In particular, when associated with a MD simulation, it accounts for conformational flexibility of the pigment-protein complex, and is able to tackle the effect of pigment-protein interactions, including mutual polarization, on site energies, inter-pigment couplings, and transition dipole strengths. The efficiency and numerical accuracy of this methodology, however, rely on a number of technical issues. Here, we provide an analysis of such issues, most notably the sensitivity of the results on the polarizable force field used to describe the MM region of the complex, as well as the effects of the

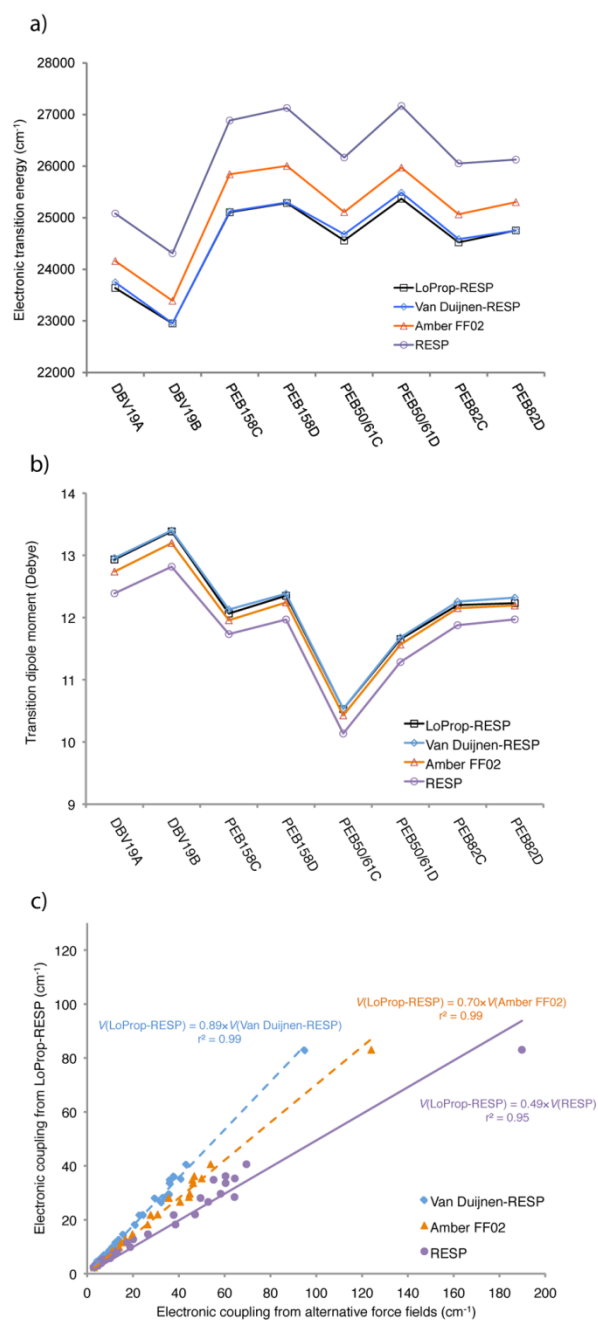
introduction of a cutoff radius to turn off MM polarization beyond a certain distance from the chromophores described at the QM level.

### **Polarization cutoff radius**

As described in the Methods section, the MMPol method solves for mutual polarization between the QM and MM regions following either an iterative procedure or adopting a matrix inversion approach. In our test calculations on PE545, we found that the iterative method is significantly faster than the matrix inversion approach. In any case, the computational cost of the calculations depends dramatically on the number of MM polarizable sites. Thus, we have introduced a cutoff radius,  $R$ , in order to turn off MM polarization in regions located beyond a certain distance from the QM region (chromophores). This implies zeroing the polarizability of any MM site located farther than  $R$  from any QM atom, while atomic charges for all MM atoms are still accounted for. In Figure S1 in the Supporting Information, we show the increase in CPU time as a function of  $R$  corresponding to a full calculation of EET parameters for one structure of the PE545 complex, i.e. the calculation of the 8 bilin excited states and the 28 electronic couplings between such states. Roughly, passing from  $R=10$  Å to 30 Å, the number of MM polarizable sites increases from  $\sim 2000$  to  $\sim 20000$ . Such increase translates into a  $\sim 6$ -fold increase in CPU time. However, up to  $R=18$  Å the variation is moderate ( $\sim 50\%$ ), whereas at larger  $R$  values the increase is more pronounced.

Obviously, the choice of an optimal  $R$  value must balance a minimal CPU time with a satisfactory degree of convergence in the electronic transition energies, transition dipole strengths and electronic couplings calculated. In Figure S1, we also show the dependence of the errors in computed EET parameters with respect to the choice of  $R$  value. Here we use as reference the  $R=30$  Å calculation, as negligible changes are already found between  $R=25$  Å and 30 Å results (see Figure S1). Electronic couplings depend on  $R$  through the explicit environment-mediated term (see Ref. [24]), while the Coulomb interaction between transition densities remains unaffected. At  $R\approx 18$  Å the differences in computed couplings are already less than 6%. This represents an error of  $\sim 1$ -2  $\text{cm}^{-1}$  for a subset of 8 chromophore pairs, and negligible ( $<1$   $\text{cm}^{-1}$ ) errors for the other 20 interactions in PE545. On the other hand, the convergence of computed transition dipole strengths is very fast, with errors  $<0.02$  Debye at  $R\approx 18$  Å, which

corresponds to changes less than 0.2%. Probably the most delicate issue concerns the determination of the site energies of each chromophore, because even small errors could lead to a different energy ordering of the bilins in the complex. In this case, at  $R \approx 18$  Å the mean error amounts to  $\sim 30$  cm<sup>-1</sup>, with the largest error (63 cm<sup>-1</sup>) found for the DBV<sub>19B</sub> molecule. We deem this error acceptable, as differences in site energies are typically much larger than 30 cm<sup>-1</sup>. Moreover, other sources of error present in the method, such as the QM/MM treatment of pigment-protein interactions, or the neglect of short-range dispersion-repulsion effects on these interactions, suggest that it is probably not necessary to pursue a finer numerical accuracy at the expense of a drastic increase in computational cost. Because of this, we adopt an  $R=18$  Å cutoff radius for the MMPol calculations performed along the MD trajectory.



**Figure 2.** Dependence of CIS/6-31G MMPol results for an individual structure of the PE545 complex extracted from the MD simulation on the classical force field used to describe the MM region. a) Site energies, b) Transition dipole moments, and c) interpigment couplings. For the latter, the correlation between results obtained with different force fields compared to the LoProp–RESP model is shown.

## Classical force field

A key issue in the application of QM/MM methods relies on the choice of the classical force field (FF) used to describe the MM region. In Figure 2, we show the comparison between the results obtained using three different FFs, as described in Section 2.2. In contrast to the QM/MM calculations performed along the MD simulation, here we use a polarization cutoff radius  $R=20$  Å in order to further minimize numerical errors in the comparison of the different FFs. As expected, the systematic underestimation of polarization effects in the Amber ff02 FF is reflected in the results. In terms of site energies, this translates into ff02 values being shifted to the blue by a roughly constant amount compared to LoProp-RESP and van Duijnen-RESP results. Even if polarization is fully turned off in the RESP force field, the relative energy differences between sites remain very similar. Interestingly, the results obtained adopting LoProp or van Duijnen parameters show an excellent agreement, despite the fundamental differences in the way the atomic polarizabilities were derived. On the other hand, transition dipole moments seem to be only slightly dependent on the polarizable FF adopted. The most remarkable impact is probably found for electronic couplings, given that the explicit environment-mediated contribution to the coupling directly depends on environment polarization. This means that if a non-polarizable FF is used, no environment screening effects are included in the estimated couplings.<sup>24</sup> Because of this, the systematic underestimation of polarization in the Amber ff02 FF is again apparent in coupling values larger by about ~30-40% compared to the other FFs, given that screening effects are significantly underestimated. The RESP results are even worse, with couplings overestimated by ~100%, given that screening effects are completely missing. In contrast, LoProp-RESP and van Duijnen-RESP couplings are very similar, with the latter being approximately 10% smaller. Overall, this analysis nicely shows that when the polarizable FF is based on MM parameters that are derived consistently with the way intramolecular polarization is handled, such as the LoProp or van Duijnen-based FFs, EET results are very robust both in terms of site energies and electronic couplings.

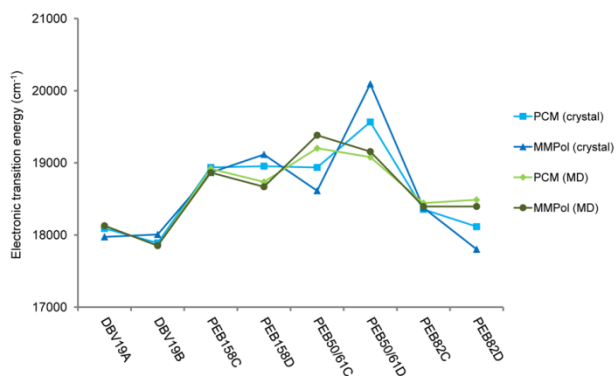
On the other hand, it is also interesting to compare the CPU time required by the different methods to compute all PE545 site energies and interpigment couplings. When the van Duijnen-RESP method is used, all mutual dipole-dipole interactions must be



computed through the Thole modified tensor, whereas the LoProp-RESP method excludes intra-residue dipole-dipole interactions. This difference translates into the LoProp-RESP method being  $\sim 2$  times faster in terms of CPU time compared to the van Duijnen-RESP method. Finally, we note that inclusion of polarization through the LoProp method increases the required time to compute all PE545 EET parameters  $\sim 2.5$  times compared to the RESP method, where only charges are included in the MM description. In the following we choose the LoProp-RESP force field to perform QM/MM calculations along the MD trajectory, as this force field shows the best balance between accuracy and computational efficiency.

### 3.2 Molecular basis for site energy tuning

In this section we analyze the molecular basis for the protein tuning of site energies. Bilins are linear tetrapyrroles with a remarkable degree of conformational flexibility. Thus, the protein can modulate site energies through local pigment-protein interactions, but also by constraining the conformational space of each individual site. For instance, a slight rotation of a bond separating the central pyrrole rings induced by the protein environment should decrease the conjugation of the bilin  $\pi$  system, leading to an increase in the electronic excitation energy. We combine atomistic QM/MM and continuum PCM calculations in order to dissect these two effects. First, we performed PCM calculations along the MD trajectory, where energy differences among DBVs, or among PEB chromophores, arise because individual sites sample different structures due to their particular location in the protein scaffold. Such continuum calculations include pigment-protein interactions, but the protein environment is described as a continuum dielectric. Then, we performed MMPol calculations in which local pigment-protein interactions are fully included in atomic detail. Here we do not discuss the impact of the heterogeneous polarizable environment on electronic couplings, as we already discussed this issue extensively on our previous study on PE545.<sup>13</sup>

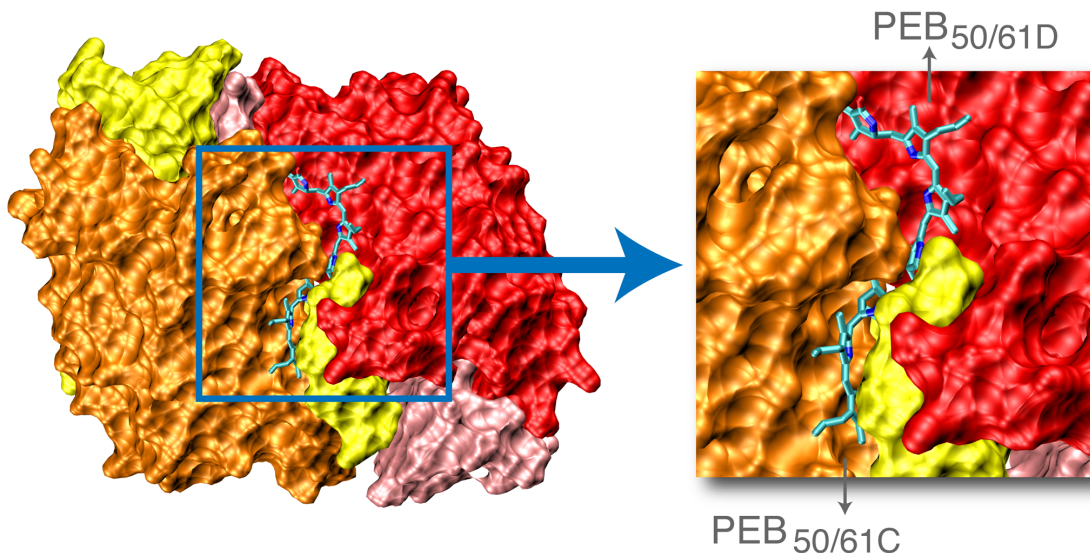


**Figure 3.** Site energies (in  $\text{cm}^{-1}$ ) predicted by PCM and MMPol calculations either based on the crystal structure or averaged over a classical trajectory of the PE545 complex (site energies and standard deviation of their values induced by structural fluctuations along the MD trajectory are reported in Table S1 in Supporting Information).

In Figure 3 we report the site energies obtained using either a continuum or an atomistic MMPol description of the protein–solvent environment. Our aim here is to assess the relative energetic ordering of the bilin sites. Given that CIS systematically overestimates transition energies due to the neglect of electron correlation, all DBV and PEB chromophores are shifted by a constant energy value, as indicated in Table S1 in the Supporting Information. Because the bilin energies are very sensitive to the particular conformation adopted by the pigment, we compare our results averaged over the MD trajectory to those obtained based on the crystal structure. A first look at the results based on the crystal structure indicates that local pigment–protein interactions do not play a major role in tuning the relative site energies, given that similar trends are obtained either from PCM or MMPol calculations. Thus, in terms of pseudosymmetric pairs, the crystal structure predicts an energetic ordering  $\text{DBV}_{19} < \text{PEB}_{82} < \text{PEB}_{158} < \text{PEB}_{50/61}$  regardless of the solvation method adopted. This suggests that the different conformations adopted by the PEBs depending on their location in the protein scaffold largely determines their energy. Thus, given that pseudosymmetric molecules, for instance  $\text{PEB}_{158\text{C}}$  and  $\text{PEB}_{158\text{D}}$ , adopt similar conformations in the protein complex, similar energies are predicted for them.

For the central  $\text{PEB}_{50/61}$  pair, however, both PCM ( $\sim 600 \text{ cm}^{-1}$ ) and MMPol methods ( $\sim 1500 \text{ cm}^{-1}$ ) predict a large energetic difference, so that the  $\text{PEB}_{50/61\text{C}}$  site has an energy

level similar or even lower than the  $\text{PEB}_{158}$  molecules. A similar energy difference is also obtained with MMPol for the  $\text{PEB}_{82}$  pair ( $\sim 600\text{cm}^{-1}$ ). The energetic differences between  $\text{PEB}_{50/61\text{C}}/\text{PEB}_{50/61\text{D}}$  or  $\text{PEB}_{82\text{C}}/\text{PEB}_{82\text{D}}$  molecules could arise from the different sequence of the  $\alpha$  polypeptide chains in the  $\alpha_1\alpha_2\beta\beta$  dimer structure of the PE545 complex, which slightly breaks the symmetry of the system, but it could also be due to the neglect of protein flexibility or inaccuracies in the crystal structure. Apart from some differences in amino acid composition, the main difference between the  $\alpha$  subunit A and B is that the A chain is 9 amino acids larger. This difference translates into an asymmetric arrangement of protein loops interacting with the  $\text{PEB}_{50/61\text{C}}/\text{PEB}_{50/61\text{D}}$  sites.<sup>23</sup> The enlarged segment of the A chain largely interacts with the  $\text{PEB}_{50/61\text{C}}$  molecule, as shown in Figure 4, whereas the  $\text{PEB}_{50/61\text{D}}$  site only slightly contacts the B subunit. This asymmetric arrangement translates into the  $\text{PEB}_{50/61\text{D}}$  site being considerably more exposed to the solvent compared to  $\text{PEB}_{50/61\text{C}}$ , and this could have direct consequences on its conformation and energy. For instance, the first solvation shell of  $\text{PEB}_{50/61\text{D}}$  contains an average of 33 waters during the MD trajectory, compared to just 20 for  $\text{PEB}_{50/61\text{C}}$ .



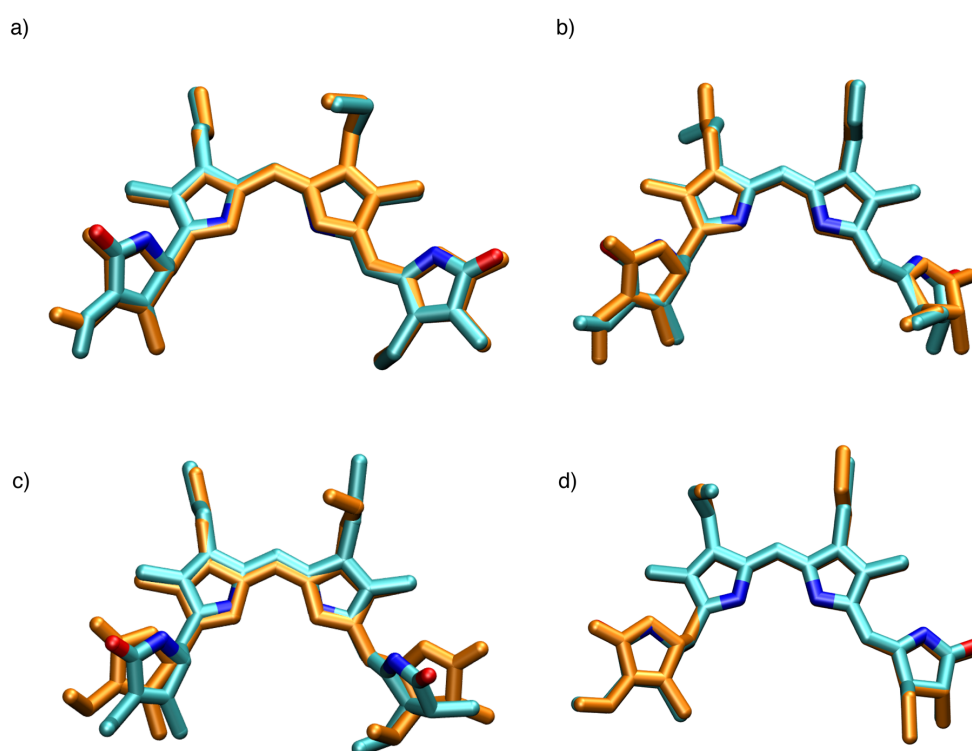
**Figure 4.** Arrangement of the  $\text{PEB}_{50/61\text{C}}$  and  $\text{PEB}_{50/61\text{D}}$  pigments in the  $\alpha_1\alpha_2\beta\beta$  structure of the PE545 complex. Protein surfaces correspond to:  $\alpha$  subunit A (yellow),  $\alpha$  subunit B (pink),  $\beta$  subunit C (orange) and  $\beta$  subunit D (red). There is an asymmetric arrangement of protein loops interacting with the pigments because the  $\text{PEB}_{50/61\text{C}}$  site closely interacts with the A and C subunits, whereas the  $\text{PEB}_{50/61\text{D}}$  site is only loosely

packed by the C and D chains, an arrangement significantly more exposed to the surrounding water solvent.

Despite the slightly asymmetric structure of PE545, when the full conformational flexibility of the pigment-protein complex is taken into account by averaging the energies over the MD trajectory, the energy difference between pseudosymmetric pigments is largely smoothed, in all cases being lower than  $300\text{ cm}^{-1}$ , in better accord with chemical intuition. Thus, the large energetic asymmetries observed for the crystal structure arise from the fundamental approximation of using a single static structure to describe the conformation of the pigment-protein complex. In addition, the energetic ordering  $\text{DBV}_{19} < \text{PEB}_{82} < \text{PEB}_{158} < \text{PEB}_{50/61}$  does not change. In this case, the agreement between MD-averaged PCM and MMPol energies is quite striking, given that the former neglect specific pigment-protein interactions believed to play a major role in tuning the site energies of other photosynthetic complexes, such as FMO.<sup>16</sup> Again, this result indicates that the protein scaffold in PE545 mainly tunes the energy ladder in the complex by fixing the individual conformation of each PEB site, rather than by specific pigment-protein interactions. Of course, additional tuning is achieved by assembling two different bilin types (DBV and PEB) in the antenna.

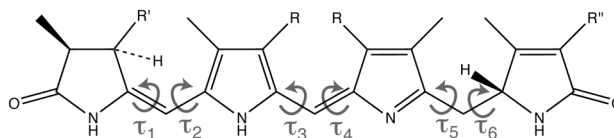
In Figure 5, we show the MD-averaged conformations of the PEB and DBV sites in PE545 that dictate the energy ladder in the complex. For clarity, pseudosymmetric molecules are superimposed, so that slight differences in their conformations can be easily observed. This picture helps us illustrating the different conformational distributions sampled by each pigment. Note, however, that our MD-averaged site energies do not exactly correspond to these conformations, as they are averaged over 141 structures extracted from the MD trajectory. As expected, the conformation of pseudosymmetric sites is very similar in each case, thus explaining their similar site energies. Small differences are observed either in the orientation of the propionate groups linked to the central pyrrole rings, or in the ethylene group linked to the lateral pyrroles. Overall, the main difference between PEB and DBV conformations in terms of the linear tetrapyrrole arrangement arises from the  $\tau_2$  torsion, as illustrated in Figure 6. The complete set of torsion angles averaged over the MD trajectory (including their standard deviations), or determined from the crystal structure, are reported in Table 1. Inspection of these results indicates that relaxing the PE545 structure through the MD

simulation induces only minor changes on the average values of the torsions compared to the crystal, with the remarkable exception of the  $\text{PEB}_{50/61\text{C}}$  site. In this case, the  $\tau_2$  value shifts by  $30^\circ$ , passing from a crystal value of  $42^\circ$  to  $68^\circ$ . In contrast,  $\text{PEB}_{50/61\text{D}}$  bilin maintains this torsion close the crystal value ( $47^\circ$ ). This effect could explain the lower energy of the  $\text{PEB}_{50/61\text{D}}$  electronic state compared to  $\text{PEB}_{50/61\text{C}}$ , given that a smaller torsion angle translates into a better conjugation along the  $\pi$  system. As discussed previously (see Figure 4), the asymmetric arrangement of protein loops interacting with the  $\text{PEB}_{50/61\text{C}}$  and  $\text{PEB}_{50/61\text{D}}$  sites most probably explains this different conformation. Overall, our results indicate that biliprotein antennae seem to have a variety of mechanisms to evolve and shape their absorption profile, such as pigment composition, conformation and specific pigment-protein interactions. Our results indicate that, in contrast to chlorophyll-based antennae, pigment composition and conformation play the major role in defining the energy ladder in the PE545 complex.



**Figure 5.** Conformations of the PE545 chromophores averaged along the MD trajectory. For each pair, the structure of the pseudosymmetric pair linked to subunit B

or D is shown superimposed in orange. a) DBV<sub>19A</sub> and DBV<sub>19B</sub>, b) PEB<sub>158C</sub> and PEB<sub>158D</sub>, c) PEB<sub>50/61C</sub> and PEB<sub>50/61D</sub>, and d) PEB<sub>82C</sub> and PEB<sub>82D</sub>.



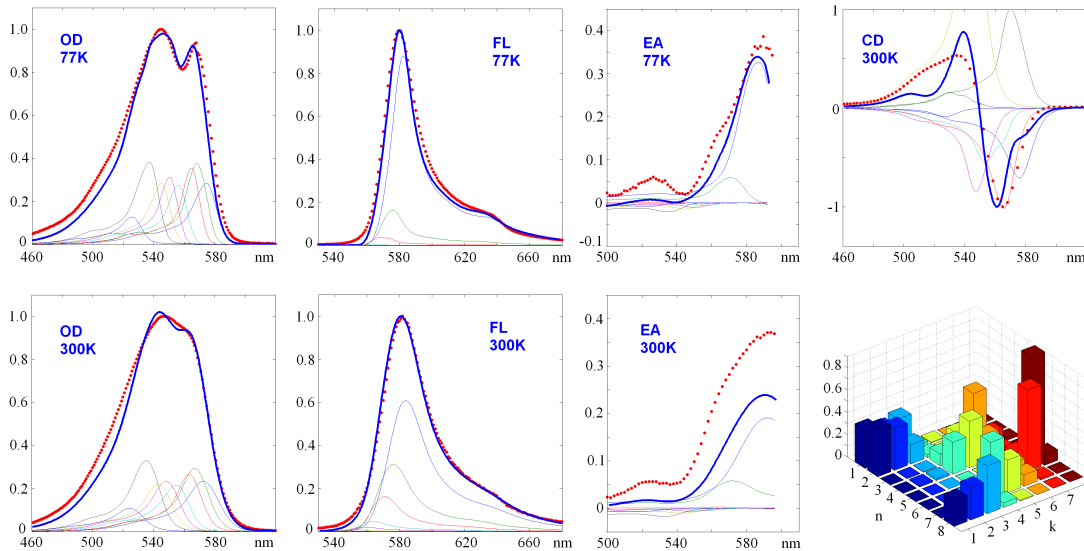
**Figure 6.** Schematic picture of the linear tetrapyrrole arrangement of bilin pigments indicating the torsions that characterize their conformation.

**Table 1.** MD-averaged dihedral angles, and their standard deviations due to structural fluctuations of the complex, characterizing the arrangement of the four pyrrole units in the pigments of PE545. For comparison, the values corresponding to the crystal structure are provided in brackets.

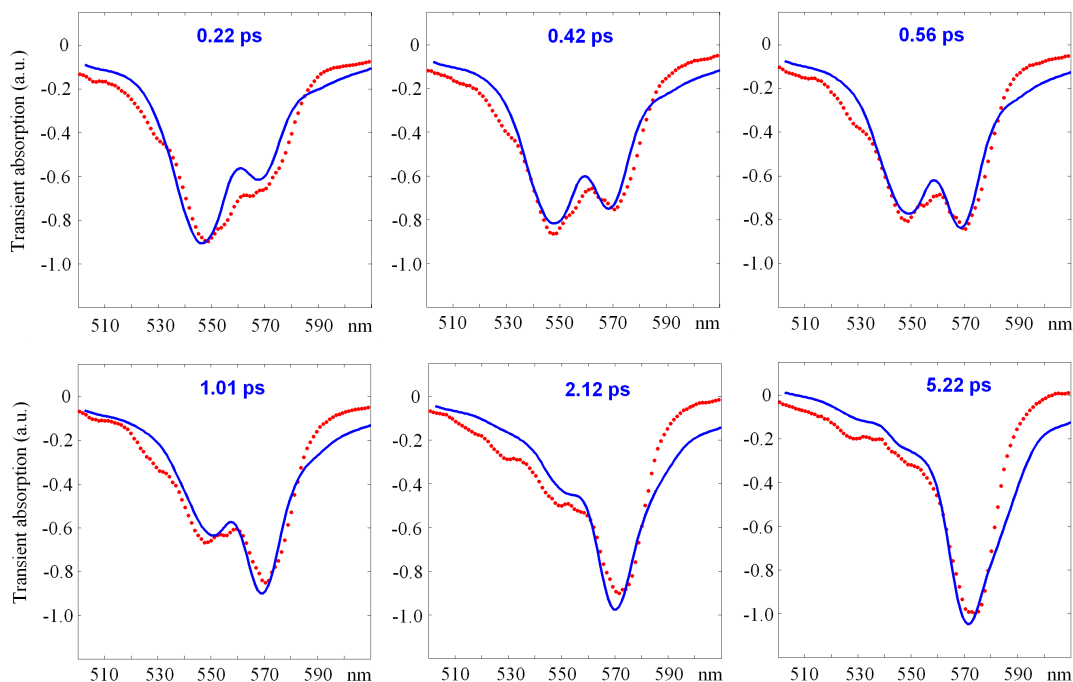
	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\tau_6$
DBV <sub>19A</sub>	183 ± 5 (182)	25 ± 57 (32)	174 ± 9 (167)	189 ± 11 (186)	250 ± 10 (255)	179 ± 7 (181)
DBV <sub>19B</sub>	185 ± 6 (184)	30 ± 80 (33)	175 ± 12 (162)	188 ± 12 (187)	252 ± 11 (251)	179 ± 6 (183)
PEB <sub>158C</sub>	181 ± 6 (176)	308 ± 10 (317)	180 ± 6 (179)	180 ± 6 (179)	260 ± 9 (268)	181 ± 6 (176)
PEB <sub>158D</sub>	178 ± 6 (175)	310 ± 9 (326)	181 ± 6 (182)	179 ± 5 (179)	263 ± 8 (267)	177 ± 6 (177)
PEB <sub>50/61C</sub>	190 ± 7 (180)	68 ± 11 (42)	189 ± 6 (184)	181 ± 5 (180)	259 ± 9 (268)	183 ± 6 (180)
PEB <sub>50/61D</sub>	182 ± 6 (181)	49 ± 8 (47)	178 ± 6 (182)	184 ± 6 (184)	272 ± 8 (278)	173 ± 7 (171)
PEB <sub>82C</sub>	185 ± 6 (179)	329 ± 10 (337)	179 ± 5 (178)	186 ± 6 (191)	265 ± 8 (269)	181 ± 7 (181)
PEB <sub>82D</sub>	188 ± 6 (183)	330 ± 10 (338)	179 ± 6 (180)	186 ± 5 (185)	262 ± 7 (272)	185 ± 7 (179)

### 3.3 Assessment of the model by comparison with spectra and kinetics

Using the site energies (Table S1), transition dipoles and center coordinates (Table S2), and couplings (Table S3) obtained with the MD-MMPol method we are able to model the steady-state spectra of PE545 (OD, CD, FL, and EA at 77K and 300K). The MMPol site energies allow reproduction of the main features of the spectra, but the positions and relative amplitudes of the 545 and 567 nm absorption subbands are slightly different from the measured ones (Figure S2 in the SI). The quality of the fit, however, can be further improved upon small adjustment of the MD-MMPol site energies, as shown in Figure 7. These adjusted site energies also give a good quantitative explanation of the TA spectra (Figure 8).



**Figure 7.** Modeling of the OD, FL, EA, and CD spectra at 77K and OD, FL, and EA spectra at 300K with the modified Redfield theory using the site energies, transition dipoles, and couplings obtained with the MD-MMPol method. The site energies have been slight adjusted as explained in the text. All the spectra are normalized (both measured and calculated) except the EA spectra, which are given in absolute units. Bar diagram shows participation of the sites  $n=1-8$  (corresponding to  $\text{DBV}_{19\text{A}}$ ,  $\text{DBV}_{19\text{B}}$ ,  $\text{PEB}_{158\text{C}}$ ,  $\text{PEB}_{158\text{D}}$ ,  $\text{PEB}_{50/61\text{C}}$ ,  $\text{PEB}_{50/61\text{D}}$ ,  $\text{PEB}_{82\text{C}}$ , and  $\text{PEB}_{82\text{D}}$  pigments) to the exciton states (from the lowest  $k=1$  to the highest  $k=8$  state). Participations are averaged over disorder ( $400\text{ cm}^{-1}$  FWHM for all the sites). Experimental spectra are the same as used in our previous modeling in Ref<sup>22</sup>.



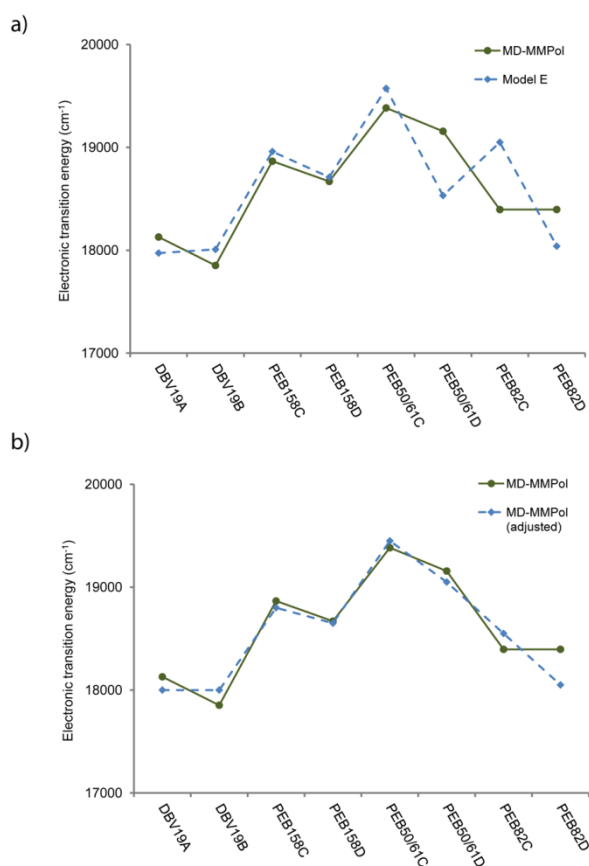
**Figure 8.** TA spectra at 77K upon 505 nm excitation at 0.22, 0.42, 0.56, 1.01, 2.12, and 5.22 ps. Measured spectra (red) are the same as in Ref <sup>22</sup>, the calculated spectra (blue) are obtained with the adjusted MD-MMPol energies.

The calculated and adjusted MD-MMPol site energies are compared with the energies extracted from the fit in our previous modeling<sup>22</sup> in Figure 9 and Table S4. Comparing the calculated MMPol energies and the adjusted MMPol energies (Figure 9b) we can find only two relatively small differences. First, the calculated  $\text{DBV}_{19\text{A}}$  and  $\text{DBV}_{19\text{B}}$  have different energies, whereas in the adjusted set this difference is eliminated in order to reproduce the red side of OD and FL shape. Second, the  $\text{PEB}_{82\text{C}}$  and  $\text{PEB}_{82\text{D}}$  pair (equal in energy in the calculation) should be more asymmetric in order to reproduce the positions, relative amplitudes, and shapes of the two absorption peaks (545 and 567 nm subbands). Thus, in the adjusted set the  $\text{PEB}_{82\text{D}}$  site is shifted to the red in order to obtain the right intensity of the 567 nm subband (together with contributions from DVB's).

The largest difference between our original model (Model E from Ref. [<sup>22</sup>]) and the current MD-MMPol calculation is the asymmetry of the  $\text{PEB}_{50/61}$  dimer. In the previous modeling it was suggested that the dimer is strongly asymmetric. Due to the large energy gap between  $\text{PEB}_{50/61\text{C}}$  and  $\text{PEB}_{50/61\text{D}}$  we found almost no coherence



between these two sites, so that the exciton levels of the  $\text{PEB}_{50/61\text{C}}\text{-PEB}_{50/61\text{D}}$  heterodimer (i.e. levels  $k=4$  and  $8$  in the original model<sup>22</sup>) are almost completely localized at one molecule, producing two spectral components with almost equal intensities. In contrast, the new MD-averaged calculations predict a much smaller gap between the  $\text{PEB}_{50/61\text{C}}$  and  $\text{PEB}_{50/61\text{D}}$  sites, i.e. the  $\text{PEB}_{50/61}$  dimer is more symmetric. In the spectra (obtained both with the calculated and adjusted energy sets) the corresponding exciton levels (i.e. levels  $k=7$  and  $8$  in Figures S2 and 7) are more delocalized. In both models the lowest level is superradiant, whereas the higher one is only weakly allowed (see the two higher exciton components of the OD spectra in Figures S2 and 7). The differences between the original model and the MD-MMPol calculation are illustrated in Figure 9a.



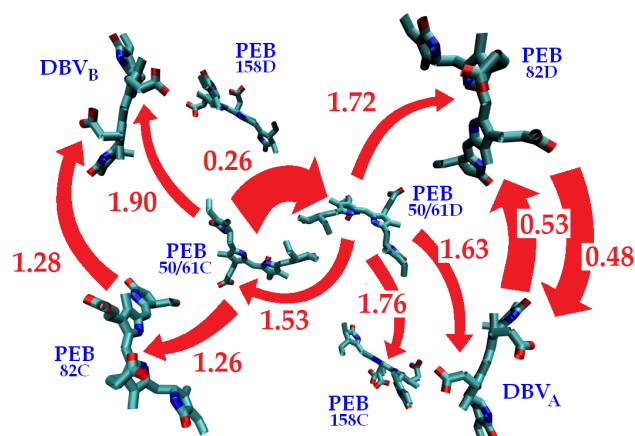
**Figure 9.** a) Site energies of the PE545 complex extracted from a fit to spectral and kinetic data (Model E in Ref <sup>22</sup>) and predicted by MD-MMPol calculations. b) Small adjustment of MD-MMPol energies in order to reproduce spectra and kinetics.

It is remarkable, that the quality of the fit obtained with the adjusted MMPol energies (as shown in Figures 7 and 8) is almost as good as obtained with the original model.<sup>22</sup> Moreover, the fit of the TA kinetics (Figure 8) is even better (the original

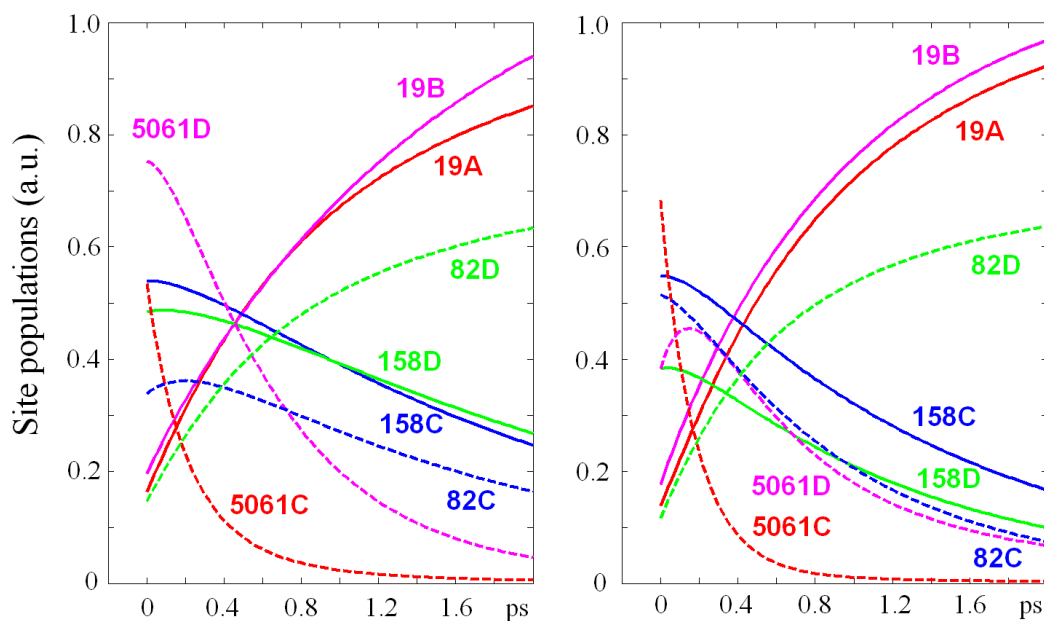
model gave a little bit too fast dynamics between 1 and 5 ps, see Figure 3 in Ref. [22]). Bearing in mind that the adjusted energies are closer to the MMPol calculation, we conclude that the current (MD-MMPol+fit) energies (Table S4) are more reliable in order to explore the spectral properties and energy transfer dynamics in the PE545 complex.

Finally, we wish to discuss the energy transfer scheme in PE545 emerging from the present modeling. The corresponding time constants, pathways of energy equilibration, and the resulting kinetics of populations are given by Table S5, Figure 10, and Figure 11, respectively. The time constants of energy transfers in the site representation are listed in Table S5. The main difference between the old (Model E) and the new scheme (based on adjusted MMPol energies) is connected with the different order in energy of the sites  $\text{PEB}_{50/61\text{D}}$  and  $\text{PEB}_{82\text{C}}$ . This changes the rates of transfer from these sites (6th and 7th columns) and to these sites (6th and 7th rows in the table). The fastest is the relaxation within the  $\text{PEB}_{50/61\text{C}}\text{-PEB}_{50/61\text{D}}$  dimer (260 fs instead of 350 fs in the old model with asymmetric dimer). Transfer from the lower dimeric site  $\text{PEB}_{50/61\text{D}}$  to the three red-most pigments  $\text{DBV}_{19\text{A}}$ ,  $\text{DBV}_{19\text{B}}$ , and  $\text{PEB}_{82\text{D}}$  is slower in the new model due to the higher position of  $\text{PEB}_{50/61\text{D}}$  energy and also to its stronger mixing with the  $\text{PEB}_{50/61\text{C}}$ . The lowest exciton state of the dimer is thus more delocalized, so transfers to the intermediate sites ( $\text{PEB}_{158\text{D,C}}$  and  $\text{PEB}_{82\text{C}}$ ) and to the red sites occur with a more uniform participation of both dimeric sites,  $\text{PEB}_{50/61\text{D}}$  and  $\text{PEB}_{50/61\text{C}}$ . The overall equilibration dynamics is expected to be slightly slower after 1 ps than in the old model.

The fastest channels of energy transfers for the new model are shown in Figure 10, where they are related to the position of the pigments in a real space.



**Figure 10.** Time constants of energy transfer (ps) in the site representation (with averaging over disorder) at 77K obtained with the adjusted MD-MMPol site energies. Transfer pathways with the time constant  $<2$  ps are shown.



**Figure 11.** Time-dependent site populations (arbitrary units) upon 505 nm excitation at 77 K calculated with the adjusted MD-MMPol site energies (left) and with the Model E from Ref <sup>22</sup> (right). Time delays are 0-2 ps. The dynamics is averaged over disorder. The sites are labeled as in Table S5.

Figure 11 shows the kinetics of the site populations upon 505 nm excitation. Notice that in the new model the lower  $\text{PEB}_{50/61\text{D}}$  dimeric site is more populated initially. This is because the vibrational wing of the lowest superradiant level is even more intense at 505 nm than the higher (symmetry-forbidden) level with predominant contribution from  $\text{PEB}_{50/61\text{C}}$ . The equilibration dynamics after 0.5-1 ps is slightly slower, as expected from the analysis of the time constants. Due to this fact we obtain better agreement between the calculated and measured TA kinetics as shown in Figure 8.

#### 4. Conclusions

In this work, we have determined the exciton model and energy transfer scheme of the PE545 antenna complex. We combined our recently developed MD-MMPol methodology with a quantitative simultaneous fit of the spectra and TA kinetics using the modified Redfield approach. The MMPol method allows for a consistent atomic detail description of electrostatic and polarization pigment-protein interactions in the estimation of site energies, transition dipole strengths and electronic couplings. In addition, conformational flexibility of the system is accounted for through classical molecular dynamics simulations.

By comparing site energies averaged over conformational disorder with predictions based on the crystal structure, we show that conformational flexibility plays a key role, which largely smoothes the large energetic differences predicted by the crystal structure between pseudosymmetric pairs. Moreover, comparison of MMPol results with analogous calculations based on the PCM continuum model, which neglects specific pigment-protein interactions, indicates that, in contrast to chlorophyll-based complexes, pigment composition and conformation play the major role in defining the energy ladder in the PE545 complex. This effect arises from the flexibility of the bilin pigments, characterized by a quasi-linear arrangement of 4 pyrrole units, thus suggesting that pigment conformation can also direct excitation energy flow in other bilin-containing photosynthetic complexes.

The MD-QM/MM site energies allow us to reproduce the main features of the spectra, and small adjustment of the energies of the three red-most pigments  $\text{DBV}_{19\text{A}}$ ,

DBV<sub>19B</sub>, and PEB<sub>82D</sub> allows us to model the spectra of PE545 with a similar quality compared to our original model,<sup>22</sup> based on a spectral and kinetic fit. Moreover, the fit of the transient absorption kinetics is even better in the new structure-based model. The largest difference between our previous and present results is that the MD-QM/MM calculations predict a much smaller gap between the pseudosymmetric PEB<sub>50/61C</sub> and PEB<sub>50/61D</sub> sites, in better accord with chemical intuition. Interestingly, this translates into a larger degree of coherence between the central PEB sites in the new model. Finally, the new results, consistent with atomistic simulations, spectra and kinetics, allow us to determine the exciton model of the complex unambiguously compared to our previous modeling, in which several sets of site energies were suggested.

## Acknowledgements

CC acknowledges support from the Ministerio de Economía y Competitividad of Spain through the Ramón y Cajal program (Ref. RYC-2011-08918). VN was supported by the visitor's grant from the Netherlands Organization for Scientific Research (NWO) (Grant No. 040.11.207) and by the Russian Foundation for Basic Research, (Grant No. 12-04-01085). JK thanks the Danish Councils for Independent Research (the STENO and Sapere Aude programmes), the Lundbeck Foundation and the Villum Foundation for financial support. AM-L acknowledges financial support from the Juan de la Cierva subprogramme of the Ministerio de Economía y Competitividad of Spain. RvG acknowledges support from the ERC via an advanced investigator grant (PHOTPROT, nr. 267333). GDS acknowledges financial support from the United States Air Force Office of Scientific Research under contract number FA9550-10-1-0260, and the Natural Sciences and Engineering Research Council of Canada. BM acknowledges the financial support from the European Research Council (ERC) (StG Enlight, n. 277755).

## Supporting Information Available

Figure reporting the dependence of MMPol results on the polarization cutoff radius, figure showing the spectral modeling obtained with the calculated MMPol site energies, tables of site energies, transition dipole moments, center coordinates of the pigments, electronic couplings and time constants of energy transfer. This information is available free of charge via the Internet at <http://pubs.acs.org>.

## References

1. van Grondelle, R.; Dekker, J. P.; Gillbro, T.; Sundstrom, V. *Biochim. Biophys. Acta, Bioenerg.* **1994**, *1187*, 1-65.
2. Novoderezhkin, V. I.; van Grondelle, R. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7352-7365.
3. McConnell, I.; Li, G. H.; Brudvig, G. W. *Chem. Biol.* **2010**, *17*, 434-447.
4. Scholes, G. D.; Fleming, G. R.; Olaya-Castro, A.; van Grondelle, R. *Nat. Chem.* **2011**, *3*, 763-774.
5. Cheng, Y. C.; Fleming, G. R. *Annu. Rev. Phys. Chem.* **2009**, *60*, 241-262.
6. Renger, T. *Photosynth. Res.* **2009**, *102*, 471-485.
7. Sener, M.; Strumpfer, J.; Hsin, J.; Chandler, D.; Scheuring, S.; Hunter, C. N.; Schulten, K. *Chemphyschem* **2011**, *12*, 518-531.
8. Mennucci, B.; Curutchet, C. *Phys. Chem. Chem. Phys.* **2011**, *13*, 11538-11550.
9. Fleming, G. R.; Scholes, G. D. *Procedia Chemistry* **2011**, *3*, 5-7.
10. Fleming, G. R.; Scholes, G. D.; Cheng, Y. C. *Procedia Chemistry* **2011**, *3*, 38-57.
11. Scholes, G. D.; Curutchet, C.; Mennucci, B.; Cammi, R.; Tomasi, J. J. *Phys. Chem. B* **2007**, *111*, 6978-6982.
12. Curutchet, C.; Scholes, G. D.; Mennucci, B.; Cammi, R. *J. Phys. Chem. B* **2007**, *111*, 13253-13265.
13. Curutchet, C.; Kongsted, J.; Muñoz-Losa, A.; Hossein-Nejad, H.; Scholes, G. D.; Mennucci, B. *J. Am. Chem. Soc.* **2011**, *133*, 3078-3084.
14. van Grondelle, R.; Novoderezhkin, V. I. *Phys. Chem. Chem. Phys.* **2006**, *8*, 793-807.
15. Adolphs, J.; Muh, F.; Madjet, M. E. A.; Busch, M. S. A.; Renger, T. *J. Am. Chem. Soc.* **2010**, *132*, 3331-3343.
16. Muh, F.; Madjet, M. E. A.; Adolphs, J.; Abdurahman, A.; Rabenstein, B.; Ishikita, H.; Knapp, E. W.; Renger, T. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16862-16867.
17. Muh, F.; Madjet, M. E. A.; Renger, T. *J. Phys. Chem. B* **2010**, *114*, 13517-13535.
18. Doust, A. B.; Wilk, K. E.; Curmi, P. M. G.; Scholes, G. D. *J. Photochem. Photobiol., A* **2006**, *184*, 1-17.
19. Olbrich, C.; Jansen, T. L. C.; Liebers, J.; Aghtar, M.; Strumpfer, J.; Schulten, K.; Knoester, J.; Kleinekathofer, U. *J. Phys. Chem. B* **2011**, *115*, 8609-8621.
20. Shim, S.; Rebentrost, P.; Valleau, S.; Aspuru-Guzik, A. *Biophys. J.* **2012**, *102*, 649-660.
21. Doust, A. B.; Marai, C. N. J.; Harrop, S. J.; Wilk, K. E.; Curmi, P. M. G.; Scholes, G. D. *J. Mol. Biol.* **2004**, *344*, 135-153.
22. Novoderezhkin, V. I.; Doust, A. B.; Curutchet, C.; Scholes, G. D.; van Grondelle, R. *Biophys. J.* **2010**, *99*, 344-352.
23. Wilk, K. E.; Harrop, S. J.; Jankova, L.; Edler, D.; Keenan, G.; Sharples, F.; Hiller, R. G.; Curmi, P. M. G. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 8901-8906.
24. Curutchet, C.; Munoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. *J. Chem. Theory Comput.* **2009**, *5*, 1838-1848.
25. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
26. Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049-1074.
27. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712-725.
28. Soderhjelm, P.; Ryde, U. *J. Phys. Chem. A* **2009**, *113*, 617-627.
29. Gagliardi, L.; Lindh, R.; Karlstrom, G. *J. Chem. Phys.* **2004**, *121*, 4494-4500.
30. Karlstrom, G.; Lindh, R.; Malmqvist, P. A.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P. O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P. et al. *Comput. Mater. Sci.* **2003**, *28*, 222-239.

31. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A. et al. *Gaussian 09, Revision A.2*, Gaussian, Inc.: Wallingford CT, 2009.
32. Case, D. A.; Darden, T. A.; Cheatham, I., T.E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M. et al. AMBER 9, University of California: San Francisco, 2006.
33. Cieplak, P.; Caldwell, J.; Kollman, P. J. *Comput. Chem.* **2001**, *22*, 1048-1057.
34. Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1995**, *99*, 6208-6219.
35. Applequist, J.; Carl, J. R.; Fung, K. K. *J. Am. Chem. Soc.* **1972**, *94*, 2952-&.
36. Ponder, J. W.; Case, D. A.; Valerie, D. *Adv. Protein Chem.* **2003**, *66*, 27-85.
37. Cieplak, P.; Dupradeau, F. Y.; Duan, Y.; Wang, J. M. *J. Phys. Condens. Matter* **2009**, *21*.
38. Thole, B. T. *Chem. Phys.* **1981**, *59*, 341-350.
39. van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399-2407.
40. Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999-3094.
41. Iozzi, M. F.; Mennucci, B.; Tomasi, J.; Cammi, R. *J. Chem. Phys.* **2004**, *120*, 7029-7040.
42. Jordanides, X. J.; Lang, M. J.; Song, X. Y.; Fleming, G. R. *J. Phys. Chem. B* **1999**, *103*, 7995-8005.
43. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024-10035.

## TOC GRAPHIC

