

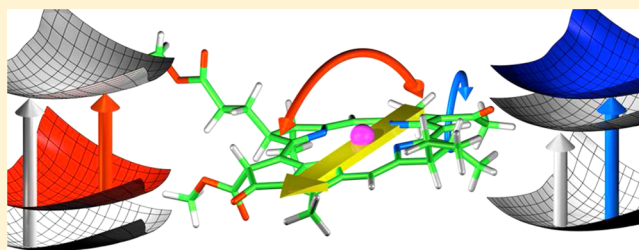
Multireference Excitation Energies for Bacteriochlorophylls A within Light Harvesting System 2

André Anda, Thorsten Hansen,* and Luca De Vico*

Department of Chemistry, Copenhagen University, Universitetsparken 5, DK-2100, Copenhagen Ø, Denmark

S Supporting Information

ABSTRACT: Light-harvesting system 2 (LH2) of purple bacteria is one of the most popular antenna complexes used to study Nature's way of collecting and channeling solar energy. The dynamics of the absorbed energy is probed by ultrafast spectroscopy. Simulation of these experiments relies on fitting a range of parameters to reproduce the spectra. Here, we present a method that can determine key parameters to chemical accuracy. These will eliminate free variables in the modeling, thus reducing the problem. Using MS-RASPT2/RASSCF calculations, we compute excitation energies and transition dipole moments of all bacteriochlorophylls in LH2. We find that the excitation energies vary among the bacteriochlorophyll monomers and that they are regulated by the curvature of the macrocycle ring and the dihedral angle of an acetyl moiety. Increasing the curvature lifts the ground state energy, which causes a red shift of the excitation energy. Increasing the torsion of the acetyl moiety raises the excited state energy, resulting in a blue shift of the excitation energy. The obtained results mark a giant leap for multiconfigurational multireference quantum chemical methods in the photochemistry of biological systems, which can prove instrumental in exposing the underlying physics of photosynthetic light-harvesting.



1. INTRODUCTION

Ninety terrawatts of solar power are absorbed by the biosphere. Plants and photosynthetic bacteria are equipped with a variety of photosystems that capture sunlight and convert it to chemical energy in the form of NADPH⁺ and ATP. The photosystems consist of light-harvesting antennas and a reaction center in which charge is separated to ultimately drive chemical reactions. The energy transfer from the antennas to the reaction center has attracted great interest due to its remarkable efficiency; in some systems, 95% of the absorbed photons successfully transfer their energy to the reaction center^{1–4} despite long distances and high noise levels.

Some antennas and reaction centers have been crystallized, and their atomic structures have been explored by X-ray diffraction.^{5–12} Different architectures of chromophore aggregates, held in place by a protein scaffold, have been revealed. The chromophores include the well-known chlorophyll and carotenoid molecules. In some families of photosynthetic bacteria, bacteriochlorophylls replace chlorophylls. This is the case for purple bacteria, which have spherical organelles covered by light-harvesting systems 1 and 2 (LH1 and LH2).

High-resolution structures of the LH1 and LH2 antenna systems have been reported in recent years.^{10–12} The structure of LH2 (a trans-membrane protein) from *Rhodoblastus acidophilus* represented in Figure 1 is highly symmetrical and consists of an outer and an inner protein cylinder made up of α -helices, which together hold two rings of bacteriochlorophyll units and carotenoids. One ring contains 9 units lying perpendicular to the α -helices, and the other ring is made up

of 18 units parallel to the α -helices. The former absorbs at approximately 800 nm and is denoted the B800 ring, whereas the close packing in the latter shifts the absorption band to around 850 nm and is denoted the B850 ring. Both rings are located in the apolar part of the protein. Closer inspection reveals that the partially overlapping units in the B850 ring are arranged alternately, like pairs of hands, depending on whether they anchor to the inner or outer protein cylinder, as shown in Figure 2. The units are denoted as BChl- α or BChl- β , respectively.

The structure alone, however, does not by itself explain the high performance of the antenna systems. To probe the dynamics of the absorbed energy, the *exciton*, ultrafast spectroscopies are being used. Processes across multiple time scales have been characterized by transient absorption^{13–15} and, in recent years, two-dimensional electronic spectroscopy (2DES).^{16–24} Spreading the molecular information out on two dimensions can potentially resolve energy transfer pathways; however, the high density of states and broad peaks, inherent in the LH2 system, result in overlapping features making it difficult to unambiguously relate spectral information to the underlying physics. A detailed understanding of the structure–function relationship requires confrontation with a theoretical model.

Simulation of 2DES is computationally expensive and only feasible for toy models, typically in the form of a Frenkel

Received: November 19, 2015

Published: January 21, 2016

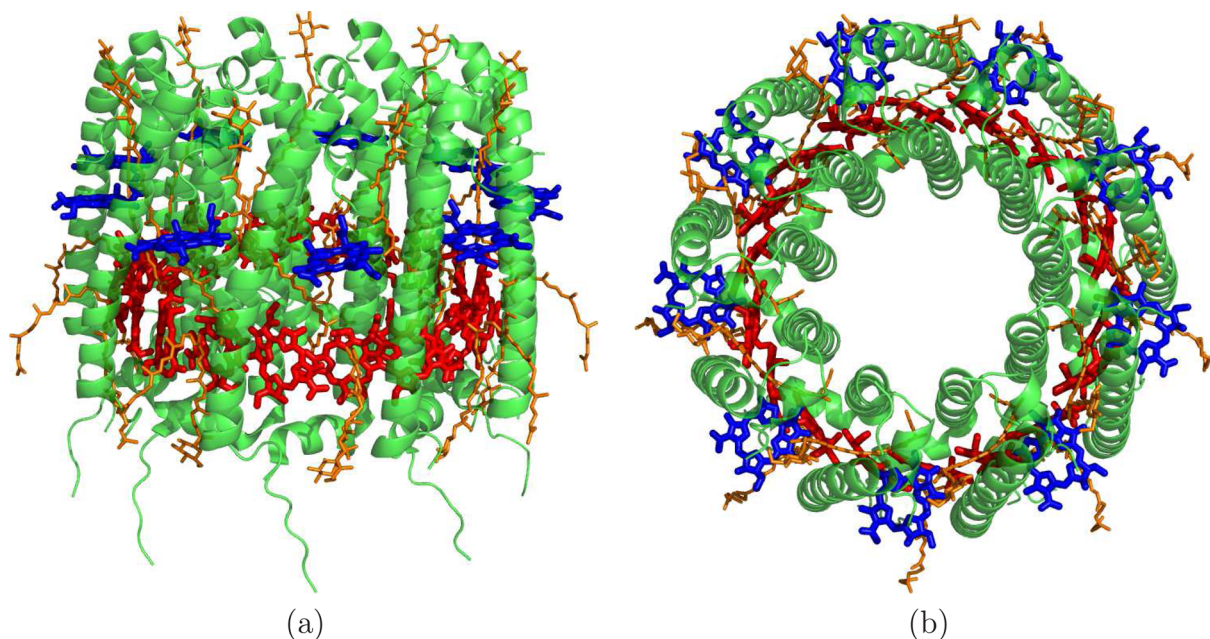


Figure 1. LH2 complex of *Rhodoblastus acidophilus*. (a) Side and (b) top view of a network of α -helices (green) that hold in place the bacteriochlorophyll units (red and blue) and the carotenoids (orange). The protein α -helices are organized in an inner and outer shell. The bacteriochlorophyll units are organized in two rings, one of 9 units absorbing at ~ 800 nm (B800, blue) and one made of 18 units absorbing at ~ 850 nm (B850, red). The bacteriochlorophyll phytol tails have been omitted for clarity.

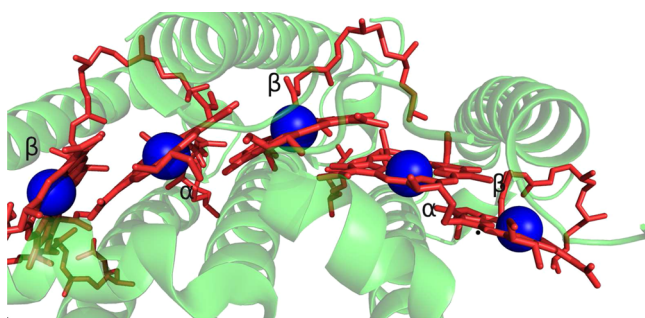


Figure 2. Bottom view of the LH2 complex. In evidence the bacteriochlorophyll units belonging to the B850 ring (in red, with the central Mg atom as blue spheres). The various units are classified as α or β depending on whether the anchoring α -helix (in green) constitutes the inner or outer shell of the protein complex.

Exciton Hamiltonian.^{25–28} Key parameters for this simplified model include excitation energies and transition dipole moments for chromophore monomers, coupling among monomers, and the spectral density describing the impact of the environment on the chromophores. To date, models of exciton dynamics in the literature are largely based on fits to a range of spectroscopic experiments. Previous attempts^{29–32} to obtain parameters ab initio demonstrates that the excitation energies of BChl units in LH2 requires³² a multireference computational method that until now has been prohibitive because of the sheer size of the chromophores. Because no experimental data is available for the excitation energies of the single BChl units embedded in the LH2 complex, it is not possible to calibrate, e.g., the TD-DFT method in terms of functionals and basis sets to obtain the desired parameters for all units.

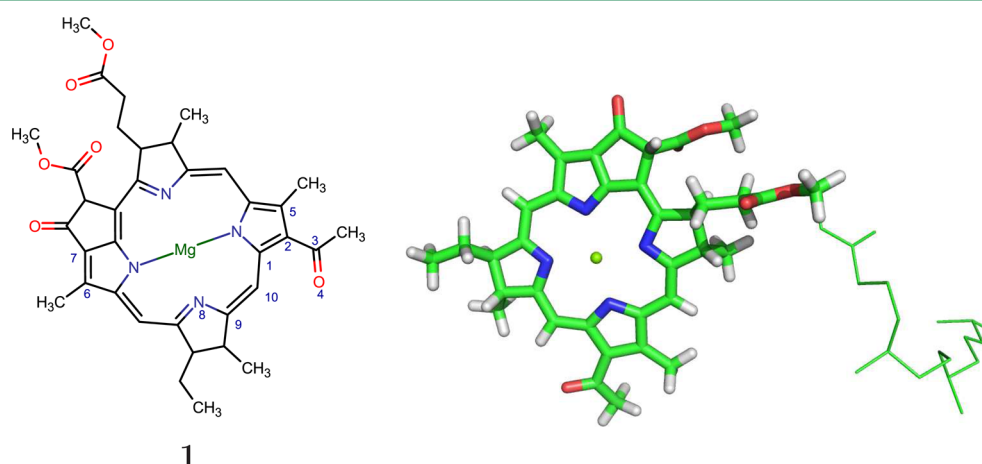


Figure 3. Employed BChl model (sticks) corresponding to **1**, where the phytol tail (lines) was removed and substituted with a methyl group.

In this work, we introduce a multiconfigurational multi-reference method capable of computing excitation energies of single units and even dimers. We present site energy calculations of unprecedented accuracy for each bacteriochlorophyll chromophore of LH2. Both site energies and transition dipole moments are obtained from MS-RASPT2/RASSCF calculations. These numbers offer a detailed view into the energetics of bacteriochlorophylls and will serve as benchmark parameters in future modeling of ultrafast spectroscopy of LH2 antenna systems.

2. COMPUTATIONAL METHODS

2.1. Preparation of Structures. The crystal structure of the LH2 complex from *Rhodoblastus acidophilus* (previously known as *Rhodospseudomonas acidophila*)³³ with access code 1NKZ was employed.¹²

The protein structure was visualized and analyzed with PyMol.³⁴ The reported LH2 complex crystal structure has a 3 fold symmetry.¹² In other words, out of the total 27 bacteriochlorophyll molecules, it contains only a subset of 9 unique BChl units, which we called units 301–309 following the original residue numbering. Six units belong to the B850 ring (units 301–306), and the remaining 3 units (307–309) belong to the B800 ring. The coordinates of the heavy atoms of the 3 α , 3 β B850 units, and 3 B800 units were extracted. Subsequently, using the program Avogadro,³⁵ the saturated phytol tail of the BChl molecule was replaced by a single carbon atom (see Figure 3), and finally, hydrogen atoms were added to obtain 1. The coordinates of all employed structures are given as Supporting Information. In the text, BChlb refers to the β B850 BChl unit that in the original crystal structure is denoted as residue 304 (site 304) after elimination of the phytol tail and optimization of the hydrogen atoms' positions. Although we are aware that BChl coordinates from the crystal structure do not necessarily represent equilibrium structures, we assume they provide a good representation of each single BChl units and, most of all, protein matrix-induced average geometrical distortion.

2.2. Energy Evaluation. For each BChl structure, we performed single point energy calculation at the SA-RASSCF/MS-RASPT2 level of theory. All reported calculations were performed on neutral molecules in vacuo. As previously noted, the bacteriochlorophyll units are embedded in the apolar part of the protein. Thus, we do not expect substantial electrostatic effects from the protein environment onto the BChl units' excitation energies, as already seen for similar systems.^{36,37} Furthermore, previous work on FMO³² showed that inclusion of the protein environment as point charges affected BChl chromophore excitation energies only if a charged amino acid was very close, which is not the case for LH2.

For each BChl unit, we computed 2 roots, state-average (SA) restricted active space self-consistent field (SA-RASSCF) single point energies. RASSCF is a more general extension of the complete active space self-consistent field method (CASSCF).³⁸ As in any MC-SCF method, a multiconfigurational wave function is constructed by linear combination of configuration state functions (CSFs). The number and type of available functions is determined by the active space (AS) and the overall spin (singlet in the present case). In RASSCF, the AS orbitals are divided into three parts, namely RAS1, RAS2, and RAS3. CSFs are built by distributing the active electrons into the active orbitals following three simple rules and maintaining the overall spin. The orbitals included in RAS2 can

have any occupation. The orbitals in RAS1 are fully occupied in all CSFs but for a number of possible holes. The orbitals in RAS3 are empty in all CSFs but for a number of possible excitations.

In the current study, we employed and tested different choices of AS to compute the energy of BChl units. BChl electronic excitations are dominated by π - π^* transitions located on the macrocycle ring (MCR). Therefore, the logical choice of orbitals to include in the AS is for the π orbitals of the MCR part. LH2 BChl units are almost planar in their MCR portion with the exception of the acetyl moiety (Figure 4 and Table 1S). The torsion around the C₂-C₃ bond is expected to regulate how much the C₃-O₄ double bond π orbitals contribute to the MCR π system.³⁹

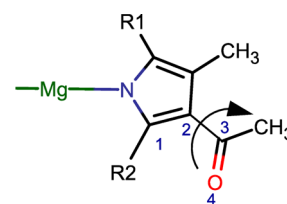


Figure 4. Acetyl moiety of 1 demonstrating the torsional angle around the C₂-C₃ bond that regulates how much the C=O π orbitals contribute to the overall BChl MCR conjugated system.

A complete description of how we chose the AS is given as Supporting Information. The final AS, denoted as AS10, comprised 26 active electrons distributed among 25 π - and π^* -type active orbitals further divided into 11 RAS1 orbitals, 4 RAS2 orbitals, and 10 RAS3 orbitals with 3 holes and excitations. The 4 RAS2 orbitals and their importance to obtain a more reliable excitation energy, as shown in the Supporting Information, are reminiscent of the four orbitals model developed a number of years ago to study the spectra of porphyrins and associated molecules.^{40–42}

The computed SA-RASSCF wave functions were used as reference for subsequent multistate, second order perturbation theory (MS-RASPT2) calculations. RASPT2⁴³ is the extension of the second order perturbation theory method CASPT2^{44,45} applied to a RASSCF-type reference wave function. MS-RASPT2 is the corresponding multistate treatment, similar to MS-CASPT2.^{46,47} Recently, MS-RASPT2 has been shown to retrieve excitations energies as good as MS-CASPT2 for different organic compounds.⁴⁸

All calculations employed an ANO-RCC basis set⁴⁹ with a double- ζ quality. The rationale behind the choice between double- and triple- ζ is given as Supporting Information. After a preliminary SCF calculation, 1s orbitals of all non-hydrogen atoms plus the 2s and 2p orbitals of magnesium were kept frozen in all subsequent RASSCF and MS-RASPT2 calculations. All single point energy calculations employed Cholesky decomposition.^{50–52} All MS-RASPT2 calculations employed standard IPEA shift⁵³ and an imaginary level shift value of 0.1.⁵⁴ MS-RASPT2 calculations were performed with 300 deleted virtual orbitals (see Supporting Information for a rationale on the choice of this number) so as to reduce the computational costs. This procedure is somewhat similar to that described by Aquilante et al.,⁵⁵ which however requires the use of quasicanonical orbitals not available for a RASSCF-type calculation.

All SA-RASSCF/MS-RASPT2 calculations were performed using MOLCAS version 7.8.⁵⁶ Specific features present only in the developer version were utilized to split the MS-RASPT2 calculations into two separate jobs, each running on a different CPU to accelerate the process. These features will soon be available in MOLCAS version 8.2.⁵⁷ Preliminary hydrogen atom position optimization was performed at the PM6⁵⁸ level of theory using the Gaussian09 suite of programs.⁵⁹

A single point energy calculation at the hybrid Coupled Cluster model CC2^{60,61} level of theory was performed using the program TURBOMOLE.⁶² It employed a cc-pVDZ basis set,^{63,64} the resolution of identity method with auxiliary basis sets from the TURBOMOLE library,⁶⁵ and the frozen core approximation.

2.3. Technical Requirements. All calculations were performed on a single processor (Intel(R) Xeon(R) CPU X5675 3.07 GHz and Intel(R) Xeon(R) CPU E5-2643 v3 3.40 GHz) and employed 25 GB of memory even if we did not test for minimum memory requirements. By using the split technique, the calculation of each MS-RASPT2 state energy using AS10 and 300 virtual orbitals required, on average, 3 weeks.

3. RESULTS AND DISCUSSION

3.1. Wave Function Analysis. Having defined the active space, we proceeded to compute MS-RASPT2/RASSCF ground and excited states energies for all BChl sites in LH2. However, given the high costs in terms of computational resources and time necessary to perform such calculations, we first provide here an analysis of the wave function obtained for BChlb, to justify the necessity of such a costly treatment.

Table 1 reports the wave function configurations with CI coefficients larger than 0.05. Raw data is reported in Table 16S. The ground state wave function is constituted by the SCF ground state configuration for ~61% ($C_0^2 = 0.60$), by single excitation configurations for at least 1%, and by double excitation configurations for at least 11% (the remaining 27% is attributed to configurations with CI coefficients smaller than 0.05). The first excited state wave function is constituted by single excitations for at least 60%, double excitations for at least 6%, and by triple excitation configurations for at least 1% with the remaining 33% due to configurations with CI coefficients smaller than 0.05.

This analysis is in accordance with previous findings for BChl obtained with DFT/MRCI.³² However, no contribution from triple excitation configurations for the first excited state has been previously reported. Moreover, the previously reported percentages of SCF ground state in Root 1 (86%) and single excitations in Root 2 (75%) were higher than what was found in the current study (~60% in both cases). Furthermore, it is interesting to note that the main single excitation configuration of the ground state is absent from the main configurations of the excited state.

To further analyze the multireference nature of BChlb and justify the usage of the MS-RASPT2/SA-RASSCF method, we performed a single point energy evaluation of BChlb at the CC2 level of theory. The obtained D1 diagnostic⁶⁶ value of ~0.10 is higher than the acceptance top value of ~0.04–0.05, as also previously found.³² The multireference character of the wave function is likely responsible for such a high D1 value.

The necessity for a multiconfigurational treatment is further exemplified by the partial electronic occupation of the active space orbitals (Table 19S). In particular, Table 2 shows how

Table 1. BChlb wave function composition for the ground (Root 1) and first excited state (Root 2), in terms of configurations, retrieved after a MS-RASPT2 calculation employing the AS10 active space^a

type	configuration	percentage
<i>Root 1</i>		
–	SCF ground state	60.5%
DE	$\pi_{\text{RAS2}}^2 \rightarrow \pi_{\text{RAS2}}^2$	5.2%
DE	$\pi_{\text{RAS2}}^1 \pi_{\text{RAS2}}^1 \rightarrow \pi_{\text{RAS2}}^1 \pi_{\text{RAS2}}^1$	3.8%
DE	$\pi_{\text{RAS1}}^1 \pi_{\text{RAS2}}^1 \rightarrow \pi_{\text{RAS2}}^1 \pi_{\text{RAS3}}^1$	2.3%
SE	$\pi_{\text{RAS1}}^1 \rightarrow \pi_{\text{RAS2}}^1$	1.0%
<i>Root 2</i>		
SE	$\pi_{\text{RAS2}}^1 \rightarrow \pi_{\text{RAS2}}^1$	59.9%
DE	$\pi_{\text{RAS1}}^1 \pi_{\text{RAS2}}^1 \rightarrow \pi_{\text{RAS2}}^1 \pi_{\text{RAS2}}^1$	2.0%
DE	$\pi_{\text{RAS2}}^2 \rightarrow \pi_{\text{RAS2}}^1 \pi_{\text{RAS3}}^1$	1.7%
DE	$\pi_{\text{RAS1}}^1 \pi_{\text{RAS2}}^1 \rightarrow \pi_{\text{RAS2}}^2$	1.3%
TE	$\pi_{\text{RAS2}}^1 \pi_{\text{RAS2}}^2 \rightarrow \pi_{\text{RAS2}}^2 \pi_{\text{RAS2}}^1$	1.2%
DE	$\pi_{\text{RAS2}}^1 \pi_{\text{RAS2}}^1 \rightarrow \pi_{\text{RAS2}}^1 \pi_{\text{RAS3}}^1$	1.0%
SE	$\pi_{\text{RAS1}}^1 \rightarrow \pi_{\text{RAS3}}^1$	0.5%

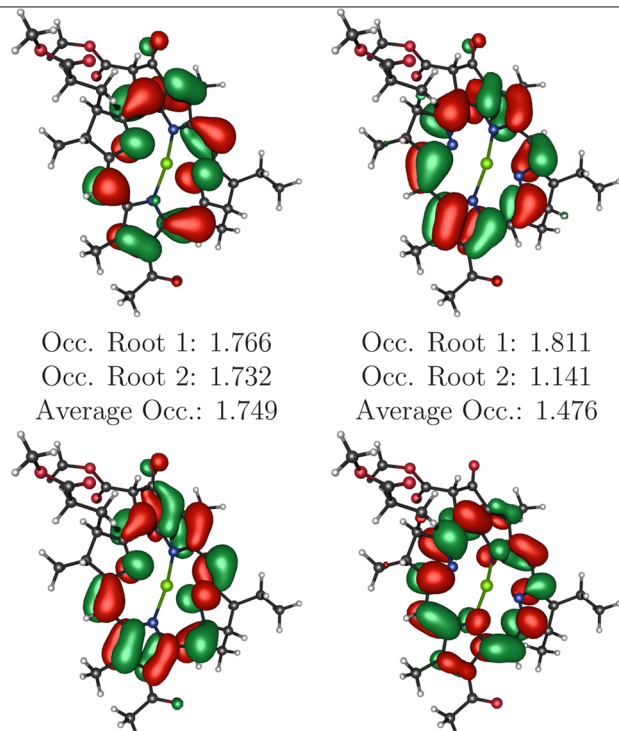
^aSingle (SE), double (DE), and triple (TE) excitations. For each configuration, type and percentage are reported. Configurations are reported in terms of which RAS sub-space the involved orbitals belong. The SCF ground state is a closed shell configuration, where RAS1 orbitals are fully occupied, only 2 of the RAS2 orbitals are fully occupied, and RAS3 orbitals are empty. The notation, e.g., $\pi_{\text{RAS1}}^1 \pi_{\text{RAS2}}^1 \rightarrow \pi_{\text{RAS2}}^1 \pi_{\text{RAS3}}^1$ indicates a configuration where one RAS1 orbital (occupied in the SCF ground state) lost 1 electron, one RAS2 orbital (occupied in the SCF ground state) lost 1 electron, one RAS2 orbital (unoccupied in the SCF ground state) gained 1 electron, and one RAS3 orbital (unoccupied in the SCF ground state) gained 1 electron. Only configurations with CI coefficients larger than 0.05 were included.

the orbitals belonging to the RAS2 subset cannot be unambiguously classified as strictly HOMO or LUMO.

Thus, the multireference character of BChlb is ascertained by three tests: (i) the SCF weight of the ground state energy $C_0^2 < 0.90$, (ii) the number of orbitals whose occupation is sensibly different from a closed shell situation, and (iii) the CC2 D1 diagnostic value higher than acceptance.

3.2. BChl Ground State Energies. Computed relative RASPT2 ground state energies are presented in Table 3. With the exclusion of units 305 and 309, BChl units belonging to the B800 ring show the lowest energies, closely followed by the α B850 BChl units. β B850 BChl units are located ~8.5 kcal/mol higher in energy than the B800 ones. This difference must stem from some geometrical factors. As previously noted (Table 15S), the different units are correlated to different torsional angles of the acetyl moiety. However, the B850 α , B850 β , and B800 units are each to be associated with a different average torsional angle, approximately 16, 23, and 29 degrees, respectively, which would not justify the ground state energy differences.

By visual inspection of the structures, it is possible to notice that the MCR of the various BChl units is curved. Such curvature can influence the overall stability of the π conjugated system. Moreover, different sites are seen as differently curved. Such curvature can be quantified from the C₂–C₆ and C₅–C₇ distances following Figure 3. Table 20S reports the retrieved distances. B800 and B850 α BChl units show, on average, similar distances. Average B850 β distances are ~0.13 Å shorter than those of the other BChl units, where shorter distances

Table 2. RASSCF computed orbitals of BChl along with their occupation^a


Occ. Root 1: 1.766 Occ. Root 2: 1.732 Average Occ.: 1.749	Occ. Root 1: 1.811 Occ. Root 2: 1.141 Average Occ.: 1.476
Occ. Root 1: 0.270 Occ. Root 2: 0.885 Average Occ.: 0.577	Occ. Root 1: 0.172 Occ. Root 2: 0.267 Average Occ.: 0.220

^aOnly the RAS2 subset of the active space is shown, as it is the most involved in the excitations reported in Table 1. The complete active space orbitals are depicted in Table 19S.

Table 3. Ground State Relative Energies of LH2 BChl Units Computed with MS-RASPT2, AS10, and 300 virtual orbitals^a

ring	site	type	relative energy	
			(kcal/mol)	(eV)
B850	301	α	1.54	0.07
B850	302	β	7.53	0.33
B850	303	α	1.39	0.06
B850	304	β	9.26	0.40
B850	305	α	5.01	0.22
B850	306	β	8.02	0.35
B800	307		0.77	0.03
B800	308		0.00	0.00
B800	309		4.90	0.21

^aThe values are relative to the lowest one, unit 308. Absolute energy values are reported in Table 22S.

mean a more curved macrocycle. Therefore, it is possible to link the relative stability of the BChl units to their curvature.

As previously noted, BChl units 305 and 309 show higher energies (~ 5 kcal/mol) than those of the same type, even if still lower than the average B850 β . An extra possible structural differentiation can be found in the relative position of the nitrogen atoms surrounding the central magnesium atom. In particular, following the numbering of Figure 3, N₈ was noticed to have different positions with respect to the quasi-plane of the

MCR. An indication of this can be obtained by considering the dihedral angle formed by the N₈-C₉-C₁₀-C₁ atoms reported in Table 21S. Unit 305 shows a different angle than the other B850 α BChl units. The same is true for unit 309 and the other two B800 units. This may represent the source of the different energies, but a further detailed analysis is out of the scope of the current study.

3.3. BChl Excitation Energies. Table 4 reports the MS-RASPT2 computed excitation energies for the various BChl

Table 4. Excitation Energies, Oscillator Strengths, and Transition Dipole Moments of LH2 BChl Units Computed with MS-RASPT2, AS10, and 300 Virtual Orbitals^a

ring	site	type	excitation energy		oscillator strength	transition dipole moment (e Å)
			(eV)	(nm)		
B850	301	α	1.66	747	0.30	1.43
B850	302	β	1.65	752	0.31	1.45
B850	303	α	1.68	737	0.30	1.44
B850	304	β	1.66	748	0.30	1.45
B850	305	α	1.68	738	0.30	1.42
B850	306	β	1.66	749	0.31	1.45
B800	307		1.69	733	0.30	1.43
B800	308		1.70	727	0.30	1.43
B800	309		1.69	734	0.29	1.41

^aAbsolute energy values are reported in Table 22S.

units of LH2 along with the corresponding oscillator strengths and transition dipole moment (TDM) sizes. The computed excitation energies fluctuate between 1.65 and 1.70 eV, corresponding to a ~ 25 nm shift. The various BChl units' excitation energies can be clearly divided according to their type. On average, B850 β units are the most red-shifted (lower excitation energy), whereas the B800 units show the largest blue shift (higher excitation energy). Apart from the energy differences, the various BChl units show very similar strength for the excitation in terms of both oscillator strength and TDM.

3.4. Geometrical Effects on Site Energies. Table 5 collates the data of Table 4 and Tables 15S and 20S to discern a possible relationship between computed excitation energy values and BChl geometrical differences. No obvious correlation can be made between average site energies and geometrical parameters. Contrary to expectations, the acetyl torsional angle is not solely responsible for tuning the excitation energies of the various BChl units. Other subtle effects from the protein matrix onto BChl structures are in place, regulating the excitation energies of each single BChl unit.

The acetyl torsion affects the excitation energy of the BChl units as long as the other parameters remain constant. This is exemplified by comparing B850 α and B800 units: to a larger torsional angle corresponds a blue-shifted excitation while maintaining a similar MCR curvature. In other words, the torsion of the acetyl moiety mainly perturbs the excited state of the BChl units because, as seen in Table 3, B850 α and B800 units have similar relative ground state energies. This is exemplified in Figure 5a.

The situation is different for B850 β units. According to the acetyl moiety dihedral angle, B850 β units should show an excitation energy average value intermediate between those of B850 α and B800 (i.e., ~ 1.68 eV). However, the computed excitation energy values are, instead, the most red-shifted. The

Table 5. Average Values of the Excitation Energies, Oscillator Strengths, and Transition Dipole Moments Reported in Table 4 Compared to Average Dihedral Angle Values of the Acetyl Moiety (Table 15S) and Average Distances between C₂–C₆ and C₅–C₇ (Table 20S) as an Indication of the MCR Curvature

ring	type	excitation energy		oscillator strength	transition dipole moment		dihedral angle (degrees)	(C ₂ –C ₆ + C ₅ –C ₇)/2 (Å)
		(eV)	(nm)		(e Å)	(debye ²)		
B850	α	1.67	741	0.30	1.43	47.2	16	8.56
B850	β	1.65	750	0.31	1.45	48.6	23	8.43
B800		1.69	731	0.30	1.42	46.5	29	8.55

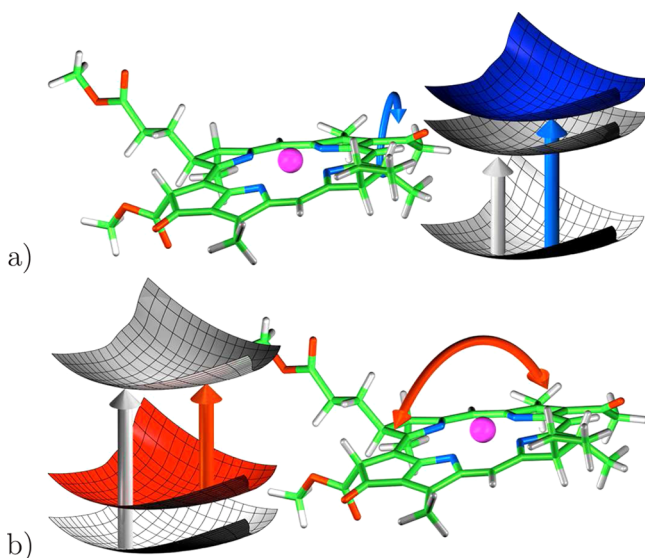


Figure 5. Effects on the computed absorption energy of geometrical changes. (a) Acetyl moiety-induced blue shift. (b) MCR curvature-induced red shift.

higher curvature of the MCR, as shown by the shorter distances reported in Table 20S and Table 5, destabilizes the ground (Table 3) more than the excited state, resulting in smaller energy differences. This is shown in Figure 5b. The effect on BChl excitation energies of the MCR curvature seems to be dominating over the acetyl torsion-based tuning.

Computed average TDM values are in line with values recorded for BChl a as extrapolated to vacuum (37.1 D²).^{67,68} The differences between the various units seem to reflect the MCR curvature. It is possible to expect that a more relaxed BChl structure (i.e., one with a lower curvature) would show a TDM value even closer to the reported extrapolated vacuum one.

3.5. Transition Dipole Moment Vectors. Figure 6 shows the BChl units and their computed TDM vectors, as immersed into the protein matrix. The TDM vectors raw data is given in Table 23S. The TDM vectors follow the general regularity of the BChl units spatial disposition. However, it is possible to note, and quantify, the relative disposition of the TDM vectors. This will be useful when simulating the spectra of the entire LH2 complex to characterize the interactions of the various TDM. The computation of accurate couplings for chromophores closer to each other less than their dimensions can be based on TDMs only as a crude approximation because the point-dipole (ideal dipole) approximation^{69,70} is not valid. However, simplified models can still be formulated for first analysis. Moreover, longer-range couplings (such as those between the B850 and B800 units) can be approximated with dipole-based approaches.

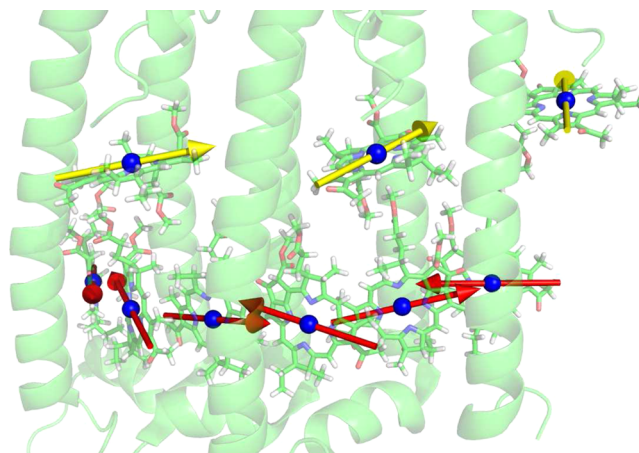


Figure 6. Representation of the computed transition dipole moment (TDM) vectors. (red) The TDM vectors relative to the B850 BChl units; (yellow) the TDM vectors of the B800 units. The vectors (Table 23S) were drawn centered on the Mg atom of each BChl unit. The vector sizes were arbitrarily adjusted for clarity.

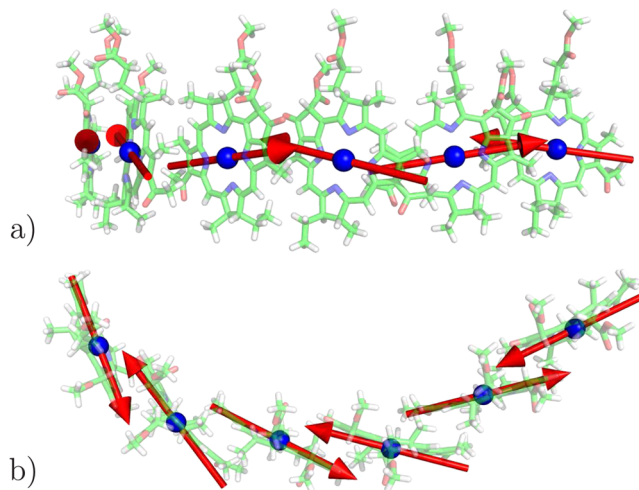


Figure 7. Side (a) and top (b) view of the TDM vectors relative to the B850 BChl units.

Figure 7 shows the TDM vectors relative to the B850 BChl units. From the side view (Figure 7a), it is noted that the TDM vectors are not all lying on the same plane but rather are following an alternating up-and-down pattern. The top view (Figure 7b) shows how the TDM vectors belonging to each α - β couple (e.g., units 301 and 302) are lying nearly antiparallel to each other, whereas those belonging to a β and the α of another couple (e.g., units 302 and 303) are clearly more rotated. This can be better appreciated by comparing the angles between the various TDM vectors, obtained as dot products, shown in Table 6. Because of the previously

mentioned up-and-down disposition, the TDM vectors belonging to, for example, units 301 and 302 are ~ 20 degrees away from being perfectly antiparallel (i.e., 180 deg). The angle between a β and an α unit belonging to the next couple (e.g., units 302 and 303) is set at ~ 145 deg, which is ~ 35 degrees away from being antiparallel.

Table 6. Angles (0–180 deg) between the Computed, Normalized TDM Vectors of the Various BChl Sites Obtained as Vector Dot Products^a

site	301	302	303	304	305	306	307	308	309
301									
302	158								
303	40	146							
304	126	39	157						
305	79	109	40	145					
306	87	79	125	40	158				
307	80	110	40	146	2	160			
308	118	71	79	109	40	147	40		
309	155	31	118	70	79	109	80	40	

^aThe upper triangle is specular to the lower one and is omitted for clarity.

The high degree of symmetry of the LH2 complex is, however, evident by considering the angle between every other BChl unit. Each α unit is rotated 40 degrees from the next α (360 deg for 9 BChl α units equal to 40 degrees per unit) and the same goes for the β units.

The symmetry of the LH2 complex is shown also by the B800 BChl units. The angle between the various B800 units (e.g., 307 and 308) is again 40 degrees. Figure 8 shows the

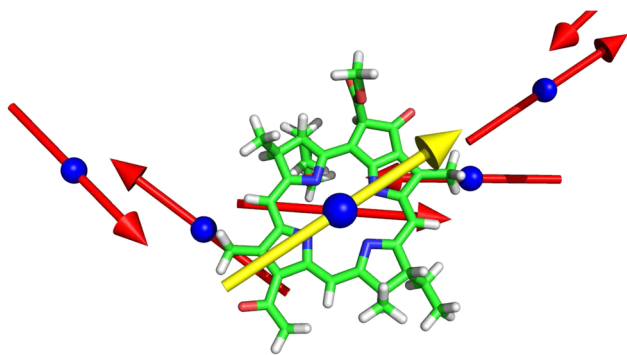


Figure 8. Top view (along the Mg atoms of units 307 and 303) of the TDM vector on unit 307 (yellow) compared to the TDM vectors (red) of the B850 units.

relative position of the TDM vector of B800 unit 307 with respect to the B850 units' vectors. The complex is shown such that the Mg atoms of units 307 and 303 are aligned. This is because 303 is the BChl unit of the B850 ring spatially closest to 307. The same goes for units 308 and 305. Unit 307 TDM has a relative angle with the TDM of unit 303 of ~ 40 degrees, whereas it is nearly colinear with the next α unit 305. The angles of the B800 BChl units TDM vectors relative to the B850 ones are expected (as partially shown in Table 6) to follow the same pattern for each single unit.

4. CONCLUSIONS

With this paper, we demonstrate that it is possible to successfully apply a multireference, multiconfigurational method, such as MS-RASPT2/SA-RASSCF, to molecules as complex (i.e., large number of heavy atoms, extended conjugated π system) as bacteriochlorophyll a. The usage of such a method is indispensable when dealing with molecules that are inherently multiconfigurational, as seen from the data reported in Table 1.

To our knowledge, the values of Table 4, together with the computed transition dipole moment vectors (Figure 6), represent the first fully multireference multiconfigurational data reported for all BChl units of the LH2 complex.

From a protein structure point of view, as shown in the discussion relative to Table 5, it is now possible to identify at least two ways with which the protein matrix fine-tunes the excitation energies of the BChl units. On one hand, the torsion of the acetyl moiety controls the first excited state of BChl and in turn can trigger a blue shift of the excitation energy. On the other hand, the overall curvature of the MCR influences the ground state of BChl and can provide a red shift to the excitation energies. It will be interesting to compare these tuning possibilities with those induced by different LH2 variants, both natural and synthetic.

Most interestingly, the observed blue shift in absorbance of LH2 complexes obtained through site-mutagenesis⁷¹ or from "low light-adapted" bacterial strains⁷² have been related to changes in the acetyl moiety torsional angle due to differences in the protein primary structures. It is our intention to relate the results of Table 5 to similar data to be obtained from BChl molecules as those contained in blue shifting complexes.

Accurate reference values for the site energies, together with the transition dipole moment vectors and their relative angles, are now available for constructing more reliable models to interpret and analyze 2DES data. The models' improvement will stem from, for example, considering the specific excitation energy of each single BChl unit and not as a whole. The reported angles between the various TDM vectors (Table 6) will also benefit the description of exciton delocalization between the various BChl molecules and the possibilities for communication between the B850 and B800 rings.

The results presented in the paper, and the methodology employed to obtain them, represent a big step forward in terms of what is currently possible to achieve with computational methods. It is our hope that the continuous development of better performing computer resources, together with the results presented here, will prompt the usage of multireference, multiconfigurational approaches for the study of photoactive molecules even more complex than BChl. We expect our data to serve as a benchmark for validating the results obtained with methods such as TD-DFT, MRCI/DFT, DMRG, or also semiempirical when applied to similar molecular systems.

Future developments and already ongoing research include the employment of the same approach as described in this paper to compute the BChl units' excitation energies to the second electronically excited state. Furthermore, we will compute the excitation energies of various complexes of the BChl units with their neighboring protein residues (e.g., Tyr 44, Trp 45, His 31) so as to introduce explicit features of the spectral density in future 2DES models. Possibly, we will extend our research also to the carotenoid molecules of LH2.

The study of the effects of the protein matrix onto BChl geometrical structures is naturally linked to the procedure of obtaining the molecules coordinates. We will investigate if and how different ways of optimizing BChl structures influence the computed excitation energies. This will include as well any eventual new, more refined crystal structure of the LH2 complex.

Finally, it is our intention to calculate accurate coupling constants between neighboring B850 BChl molecules through the use of multireference multiconfigurational techniques as those described in this paper to account for orbital overlapping of vicinal units. We will compare the obtained data with those attained through simpler dipole-based approximations,⁷⁰ transition point charge descriptions,⁷³ or transition density cubes.⁷⁴

■ ASSOCIATED CONTENT

📄 Supporting Information

This material is available free of charge via the Internet at <http://pubs.acs.org/>. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b01104.

Discussions on the choice of active space employed in this study, on the number of deleted virtual orbitals, basis set size, and the effects of RASSCF energy thresholds along with tables of the coordinates of all structures considered in this work, including geometrical parameters extracted from these coordinates, wave function composition raw data, orbitals depiction, transition dipole moment components, and complete computed energetics in absolute and relative values (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: thorsten@chem.ku.dk

*E-mail: luca@chem.ku.dk

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Jacob Kongsted is acknowledged for fruitful discussions on BChl. Per Åke Malmqvist is acknowledged for kind help in computing and visualizing the transition dipole moment vectors. T.H. is grateful for financial support from the Lundbeck Foundation. All calculations were conducted using the resources available at the Danish Center for Scientific Computing at Copenhagen University. The Center for Exploitation of Solar Energy is acknowledged for providing computational resources.

■ REFERENCES

- (1) Hu, X.; Ritz, T.; Damjanovic, A.; Autenrieth, F.; Schulten, K. Q. *Rev. Biophys.* **2002**, *35*, 1–62.
- (2) Pullerits, T.; Sundström, V. *Acc. Chem. Res.* **1996**, *29*, 381–389.
- (3) Fleming, G. R.; van Grondelle, R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 738–748.
- (4) Ferretti, M.; Duquesne, K.; Sturgis, J. N.; van Grondelle, R. *Phys. Chem. Chem. Phys.* **2014**, *16*, 26059–26066.
- (5) Camara-Artigas, A.; Blankenship, R. E.; Allen, J. P. *Photosynth. Res.* **2003**, *75*, 49–55.
- (6) Cogdell, R. J.; Isaacs, N. W.; Freer, A. A.; Arrelano, J.; Howard, T. D.; Papiz, M. Z.; Hawthornthwaite-Lawless, A. M.; Prince, S. *Prog. Biophys. Mol. Biol.* **1997**, *68*, 1–27.

- (7) Fenna, R. E.; Matthews, B. W. *Nature* **1975**, *258*, 573–577.
- (8) Li, Y.-F.; Zhou, W.; Blankenship, R. E.; Allen, J. P. *J. Mol. Biol.* **1997**, *271*, 456–471.
- (9) McDermott, G.; Prince, S. M.; Freer, A. A.; Hawthornthwaite-Lawless, A. M.; Papiz, M. Z.; Cogdell, R. J.; Isaacs, N. W. *Nature* **1995**, *374*, 517–521.
- (10) Niwa, S.; Yu, L.-J.; Takeda, K.; Hirano, Y.; Kawakami, T.; Wang-Otomo, Z.-Y.; Miki, K. *Nature* **2014**, *508*, 228–232.
- (11) Roszak, A. W.; Howard, T. D.; Southall, J.; Gardiner, A. T.; Law, C. J.; Isaacs, N. W.; Cogdell, R. J. *Science* **2003**, *302*, 1969–1972.
- (12) Papiz, M. Z.; Prince, S. M.; Howard, T.; Cogdell, R. J.; Isaacs, N. W. *J. Mol. Biol.* **2003**, *326*, 1523–1538.
- (13) Trautman, J. K.; Shreve, A. P.; Violette, C. A.; Frank, H. A.; Owens, T. G.; Albrecht, A. C. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 215–219.
- (14) Shreve, A. P.; Trautman, J. K.; Frank, H. A.; Owens, T. G.; Albrecht, A. C. *Biochim. Biophys. Acta, Bioenerg.* **1991**, *1058*, 280–288.
- (15) Herek, J. L.; Fraser, N. J.; Pullerits, T.; Martinsson, P.; Polívka, T.; Scheer, H.; Cogdell, R. J.; Sundström, V. *Biophys. J.* **2000**, *78*, 2590–2596.
- (16) Zigmantas, D.; Read, E. L.; Mančal, T.; Brixner, T.; Gardiner, A. T.; Cogdell, R. J.; Fleming, G. R. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12672–12677.
- (17) Harel, E.; Long, P. D.; Engel, G. S. *Opt. Lett.* **2011**, *36*, 1665–1667.
- (18) Harel, E.; Engel, G. S. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 706–711.
- (19) Fidler, A. F.; Singh, V. P.; Long, P. D.; Dahlberg, P. D.; Engel, G. S. *J. Phys. Chem. Lett.* **2013**, *4*, 1404–1409.
- (20) Fidler, A. F.; Singh, V. P.; Long, P. D.; Dahlberg, P. D.; Engel, G. S. *J. Chem. Phys.* **2013**, *139*, 155101.
- (21) Dahlberg, P. D.; Fidler, A. F.; Caram, J. R.; Long, P. D.; Engel, G. S. *J. Phys. Chem. Lett.* **2013**, *4*, 3636–3640.
- (22) Ostroumov, E. E.; Mulvaney, R. M.; Anna, J. M.; Cogdell, R. J.; Scholes, G. D. *J. Phys. Chem. B* **2013**, *117*, 11349–11362.
- (23) Ostroumov, E. E.; Mulvaney, R. M.; Cogdell, R. J.; Scholes, G. D. *Science* **2013**, *340*, 52–56.
- (24) Fidler, A. F.; Singh, V. P.; Long, P. D.; Dahlberg, P. D.; Engel, G. S. *Nat. Commun.* **2014**, *5*, 3286.
- (25) Renger, T.; Madjet, M.-A.; Schmidt am Busch, M.; Adolphs, J.; Müh, F. *Photosynth. Res.* **2013**, *116*, 367–388.
- (26) Renger, T.; Müh, F. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3348–3371.
- (27) Rancova, O.; Abramavicius, D. *J. Phys. Chem. B* **2014**, *118*, 7533–7540.
- (28) van der Vegte, C. P.; Prajapati, J. D.; Kleinekathöfer, U.; Knoester, J.; Jansen, T. L. C. *J. Phys. Chem. B* **2015**, *119*, 1302–1313.
- (29) Sauer, K.; Cogdell, R. J.; Prince, S. M.; Freer, A.; Isaacs, N. W.; Scheer, H. *Photochem. Photobiol.* **1996**, *64*, 564–576.
- (30) Scholes, G. D.; Gould, I. R.; Cogdell, R. J.; Fleming, G. R. *J. Phys. Chem. B* **1999**, *103*, 2543–2553.
- (31) Linnanto, J.; Korppi-Tommola, J. *J. Phys. Chem. A* **2001**, *105*, 3855–3866.
- (32) List, N. H.; Curutchet, C.; Knecht, S.; Mennucci, B.; Kongsted, J. *J. Chem. Theory Comput.* **2013**, *9*, 4928–4938.
- (33) Imhoff, J. F. *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 1863–6.
- (34) Schrödinger, LLC, *The PyMOL Molecular Graphics System*, version 1.2r1, 2010.
- (35) Hanwell, M.; Curtis, D.; Lonie, D.; Vandermeersch, T.; Zurek, E.; Hutchison, G. *J. Cheminf.* **2012**, *4*, 17.
- (36) Lammich, L.; Åxman Petersen, M.; Brønsted Nielsen, M.; Andersen, L. *Biophys. J.* **2007**, *92*, 201–207.
- (37) Milne, B. F.; Toker, Y.; Rubio, A.; Brønsted Nielsen, S. *Angew. Chem., Int. Ed.* **2015**, *54*, 2170–2173.
- (38) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. *Chem. Phys.* **1980**, *48*, 157–173.
- (39) Cogdell, R.; Howard, T.; Isaacs, N.; McLuskey, K.; Gardiner, A. *Photosynth. Res.* **2002**, *74*, 135–141.

- (40) Gouterman, M.; Wagnière, G. H.; Snyder, L. C. *J. Mol. Spectrosc.* **1963**, *11*, 108–127.
- (41) Weiss, C. J. *J. Mol. Spectrosc.* **1972**, *44*, 37–80.
- (42) Hanson, L. K. *Photochem. Photobiol.* **1988**, *47*, 903–921.
- (43) Malmqvist, P.-Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. *J. Chem. Phys.* **2008**, *128*, 204109.
- (44) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
- (45) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
- (46) Roos, B.; Andersson, K.; Fülischer, M.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. In *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*; Prigogine, I., Rice, S., Eds.; John Wiley & Sons: New York, 1996; pp 219–332.
- (47) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299–306.
- (48) Sauri, V.; Serrano-Andrés, L.; Shahi, A. R. M.; Gagliardi, L.; Vancoillie, S.; Pierloot, K. *J. Chem. Theory Comput.* **2011**, *7*, 153–168.
- (49) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.
- (50) Aquilante, F.; Pedersen, T. B.; Lindh, R. *J. Chem. Phys.* **2007**, *126*, 194106.
- (51) Aquilante, F.; Pedersen, T. B.; Lindh, R.; Roos, B. O.; Sánchez de Merás, A.; Koch, H. *J. Chem. Phys.* **2008**, *129*, 024113.
- (52) Aquilante, F.; Malmqvist, P.-Å.; Pedersen, T. B.; Ghosh, A.; Roos, B. O. *J. Chem. Theory Comput.* **2008**, *4*, 694–702.
- (53) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **2004**, *396*, 142–149.
- (54) Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196–204.
- (55) Aquilante, F.; Todorova, T. K.; Gagliardi, L.; Pedersen, T. B.; Roos, B. O. *J. Chem. Phys.* **2009**, *131*, 034113.
- (56) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Neogrády, P.; Bondo Pedersen, T.; Pitoňák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224–247.
- (57) Aquilante, F.; Autschbach, J.; Carlson, R. K.; Chibotaru, L. F.; Delcey, M. G.; DeVico, L.; Fernández Galván, I.; Ferré, N.; Frutos, L. M.; Gagliardi, L.; Garavelli, M.; Giussani, A.; Hoyer, C. E.; LiManni, G.; Lischka, H.; Ma, D.; Malmqvist, P.-Å.; Müller, T.; Nenov, A.; Olivucci, M.; Bondo Pedersen, T.; Peng, D.; Plasser, F.; Pritchard, B.; Reiher, M.; Rivalta, I.; Schapiro, I.; Segarra-Martí, J.; Stenrup, M.; Truhlar, D. G.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Vysotskiy, V. P.; Weingart, O.; Zapata, F.; Lindh, R. *J. Comput. Chem.* **2016**, *37*, 506–541.
- (58) Stewart, J. J. *Mol. Model.* **2007**, *13*, 1173–1213.
- (59) 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.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
- (60) Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409–418.
- (61) Hättig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154–5161.
- (62) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (63) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (64) Woon, D. E.; Dunning, T. H., Jr. To be published.
- (65) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.
- (66) Janssen, C. L.; Nielsen, I. M. *Chem. Phys. Lett.* **1998**, *290*, 423–430.
- (67) Knox, R. S. *Photochem. Photobiol.* **2003**, *77*, 492–496.
- (68) Knox, R. S.; Spring, B. Q. *Photochem. Photobiol.* **2003**, *77*, 497–501.
- (69) Förster, T. *Ann. Phys.* **1948**, *437*, 55–75.
- (70) Renger, T.; Madjet, M. E.-A.; Schmidt am Busch, M.; Adolphs, J.; Müh, F. *Photosynth. Res.* **2013**, *116*, 367–388.
- (71) Fowler, G. J. S.; Visschers, R. W.; Grief, G. G.; van Grondelle, R.; Hunter, C. N. *Nature* **1992**, *355*, 848–850.
- (72) McLuskey, K.; Prince, S. M.; Cogdell, R. J.; Isaacs, N. W. *Biochemistry* **2001**, *40*, 8783–8789.
- (73) Steinmann, C.; Kongsted, J. *J. Chem. Theory Comput.* **2015**, *11*, 4283–4293.
- (74) Krueger, B. P.; Scholes, G. D.; Fleming, G. R. *J. Phys. Chem. B* **1998**, *102*, 5378–5386.