Supplementary Information:

Exciton properties and optical spectra of Light Harvesting Complex II from a fully atomistic description

Vladislav Sláma,* Lorenzo Cupellini, and Benedetta Mennucci*
Department of Chemistry and Industrial Chemistry, University of Pisa, via G. Moruzzi 13, 56124 Pisa, Italy
* *E-mail: vladislav.slama@dcci.unipi.it; benedetta.mennucci@unipi.it*

S1 Spectral density dependence

The dependence of the optical spectra with the spectral density is here investigated comparing the results obtained with the function proposed by Novoderezhkin et al.[1] (SD1), with that proposed by Müh et al.[2] (SD2).

SD1

We distinguish between high-frequency pigment modes and low-frequency protein or solvent modes. The former are essentially in the vibrational ground state at low temperatures, when the pigment protein complex is in the electronic ground state. The bath correlation function of both contributions reads as

$$J(\omega) = 2\frac{\lambda_0}{\tau_c}\frac{\omega}{\omega^2 + (1/\tau_c)^2} + \sum_{i=1}^{48} 2\lambda_i \omega_i^2 \frac{\omega\gamma_i}{\left(\omega_i^2 - \omega^2\right)^2 + \omega^2\gamma_j^2}$$

$$\lambda_i = S_i\omega_i$$

where S_i correspond to Huang-Rhys factors and λ_i to reorganization energies. The parameters of the correlation function were obtained by fitting the fluorescence line narrowing spectra. The low frequency modes are represented by the over-damped Brownian oscillator spectral density with correlation time τ_c =177 fs and reorganization energy λ =37 cm^{-1}. The high frequency part of the spectral density was represented by the sum of the underdamped Brownian oscillators with parameters presented in Table S1 and inverse correlation time $\gamma = 3.0cm^{-1}$.

$\omega_i[cm^{-1}]$	S_i	$\omega_i[cm^{-1}]$	S_i	$\omega_i[cm^{-1}]$	S_i	$\omega_i[cm^{-1}]$	S_i
97	0.02396	604	0.00194	1143	0.04094	1354	0.00576
138	0.02881	700	0.00197	1181	0.01759	1382	0.00667
213	0.03002	722	0.00394	1190	0.00667	1439	0.00667
260	0.02669	742	0.03942	1208	0.0185	1487	0.00788
298	0.02669	752	0.02578	1216	0.01759	1524	0.00636
342	0.06035	795	0.00485	1235	0.00697	1537	0.02183
388	0.02487	916	0.02123	1252	0.00636	1553	0.00909
425	0.01486	986	0.01031	1260	0.00636	1573	0.00454
518	0.03942	995	0.02274	1286	0.00454	1580	0.00454
546	0.00269	1052	0.01213	1304	0.00576	1612	0.00454
573	0.00849	1069	0.00636	1322	0.03032	1645	0.00363
585	0.00303	1110	0.01122	1338	0.00394	1673	0.00097

Table S1 High frequency part of the spectral density from Novoderezhkin et al.

SD2

In the present model of the spectral density, the high-frequency modes are neglected. As a consequence, the spectral density $J(\omega)$ includes only low-frequency modes and can be described by the analytic function:[2]

$$J(\omega) = \pi S_0 \frac{1}{s_1 + s_2} \sum_{i=1,2} \frac{s_i}{7!2\omega_i^4} \omega^5 e^{-\sqrt{|\omega|/\omega_i}}$$

where the individual parameters are s_1=0.8, s_2=0.5, $\hbar\omega_1$=0.56 cm^{-1}, and $\hbar\omega_2$=1.94 cm^{-1}. S_0 is the total Huang-Rhys factor of the low-frequency modes and it was estimated from the temperature dependence of the absorbance spectrum to be $\sim$0.5.

Comparison of Optical spectra

We compared two sets of the optical spectra, Fig. S1. The first set (presented in the main text) was calculated with a spectral density which includes the coupling of all the intramolecular vibrations to the electronic transitions (SD1). The second set was calculated with a spectral density which only includes low frequency modes (SD2). These low frequency modes correspond to the coupling of the electronic transitions to solvent and protein vibrations. Contrary to the spectra of the single realization, the average spectra over the static disorder are for both spectral densities almost identical. This is due to the fact that the intrapigment (high-frequency) normal modes have relatively small Franck-Condon factors in the case of Chl Q_y transitions and therefore small oscillator strengths.

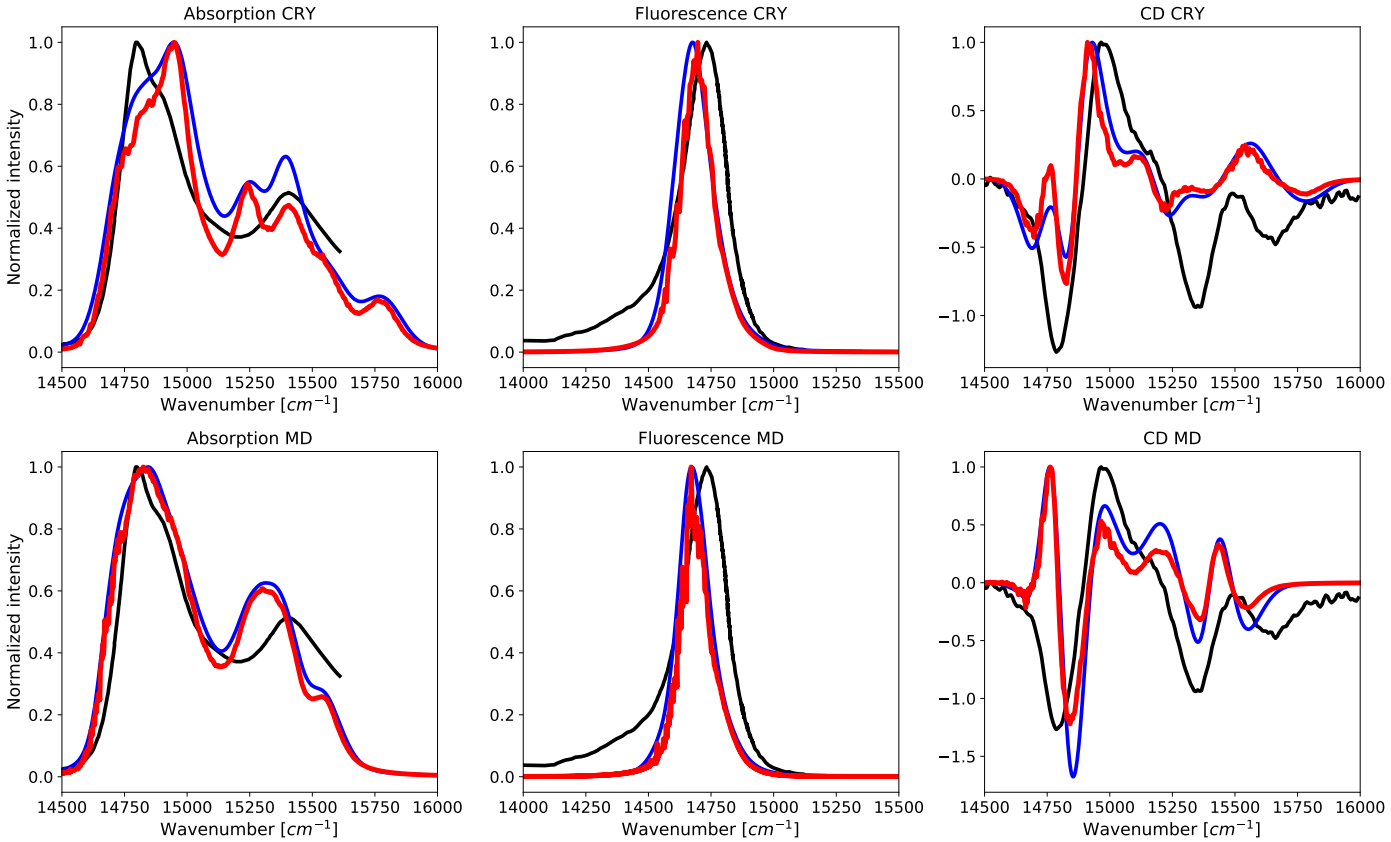


Fig. S1 Comparison of the optical spectra at 77K obtained for the different spectral densities. Blue spectra calculated with SD1 and the red ones were calculated with SD2. The black line correspond to the experimental absorption [3], fluorescence and CD spectra [4] at 77K.

S2 Effect of the lifetime broadening

To investigate effect of the lifetime broadening on the optical spectra, we compared spectra where the lifetimes of the exciton states were calculated from the Redfield rates (blue) and spectra without the lifetime broadening (red), Fig. S2. The experimental spectra are between the simulated spectra with and without the lifetime broadening. In fact, the red spectra better capture the main Chl-a peak width (for 77K MD spectra). Moreover, for room temperature, the red spectra show better agreement with the experimental fluorescence and CD spectra and the rise of the first absorption peak is described better, even though the width of the absorption spectra is slightly underestimated. All these improvements in the spectra without the lifetime broadening even if small seem to indicate an overestimation of the Redfield rates and corresponding lifetimes of the excited states.

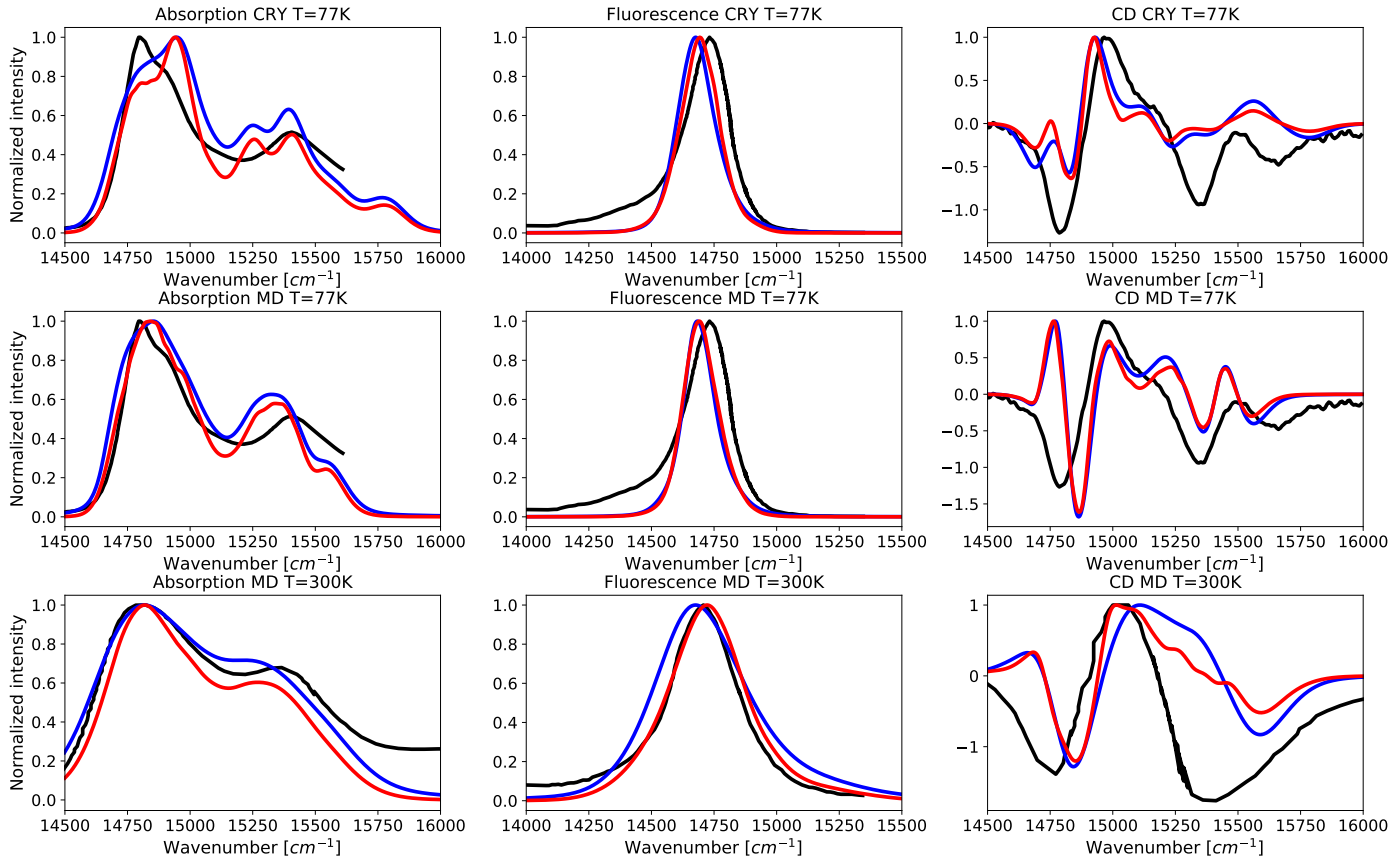


Fig. S2 Absorption, fluorescence and CD spectra calculated for crystal structure and for MD snapshots. The blue spectra were calculated with the lifetime broadening and the red one without the lifetime broadening. The black spectra correspond to the experimental absorption [3], fluorescence and CD spectra[4] at 77K and absorption[5], fluorescence[4] and CD[6] spectra at the room temperature, respectively.

S3 Excitonic properties

Table S2 Site energies in cm^{-1} calculated within TDDFT/MMpol approach for the LHCII crystal structure using different DFT functionals. The ONIOM results corresponds to calculations where the geometries of the Chls have been optimized at B3LYP/6-31+G(d) level and the site energies are calculated with M062X. All the energies are in cm^{-1}.

	M062X			ωB97XD			CAM-B3LYP			LC-BLYP			B3LYP			ONIOM all
	mon1	mon2	mon3	mon1	mon2	mon3	mon1	mon2	mon3	mon1	mon2	mon3	mon1	mon2	mon3	mon1
a602	16030	16005	16136	15298	15263	15404	15554	15518	15655	15335	15280	15434	15753	15700	15859	16813
a603	16054	16317	16077	15350	15523	15383	15592	15765	15623	15346	15549	15387	15798	16063	15821	16931
a604	16189	16141	16285	15459	15432	15575	15709	15677	15818	15524	15484	15627	16030	15938	16133	16943
a610	16233	16180	16050	15467	15412	15267	15723	15670	15533	15541	14979	15360	16071	16803	15908	16621
a611	16100	16021	16138	15314	15236	15365	15583	15510	15623	15416	15321	15447	16067	15946	16073	16913
a612	16370	16354	16330	15650	15624	15813	15871	15856	15840	15698	15679	15660	16171	16139	16055	17073
a613	16234	16402	16104	15467	15704	15317	15732	15956	15601	15584	15791	15452	16157	16021	16036	16996
a614	16264	16236	16276	15492	15454	15493	15760	15729	15766	15592	15568	15607	16240	16217	16260	16927
b601	16864	16722	16795	16229	16075	16172	16433	16288	16371	16171	16016	16085	16370	16276	16242	17688
b605	16839	16748	16858	16234	16126	16252	16416	16318	16430	16161	16071	16175	16246	16269	16309	17751
b606	16468	16443	16460	15777	15821	15767	15998	16021	16118	15715	15744	15733	15936	15833	16030	17501
b607	16545	16793	16446	15877	16139	15794	16084	16342	16003	15782	16038	15712	15922	16183	15894	17647
b608	16355	16395	16409	15740	15798	15764	15927	15978	15963	15708	15762	15764	15950	15986	16104	16932
b609	17014	17094	16842	16432	16529	16254	16636	16684	16443	16275	16368	16092	16239	16351	16071	17911

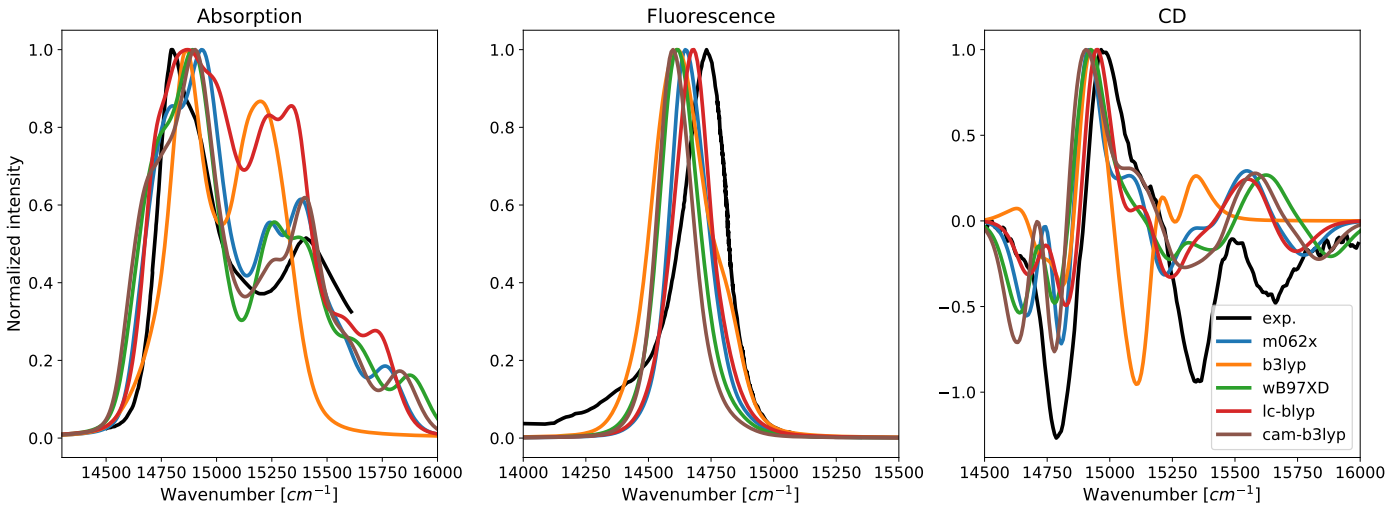


Fig. S3 TDDFT/MMPol spectra (at 77K) calculated for the crystal structure with different DFT functionals. All the spectra have been obtained by shifting the site energies shifted so to match the dominant absorption band in the experimental spectra; namely -1220 cm^{-1} for the M06-2X functional, -520 cm^{-1} for the ωB97XD functional, -760 cm^{-1} for the CAM-B3LYP functional, -500 cm^{-1} for the LC-BLYP functional and -1000 cm^{-1} for the B3LYP functional.

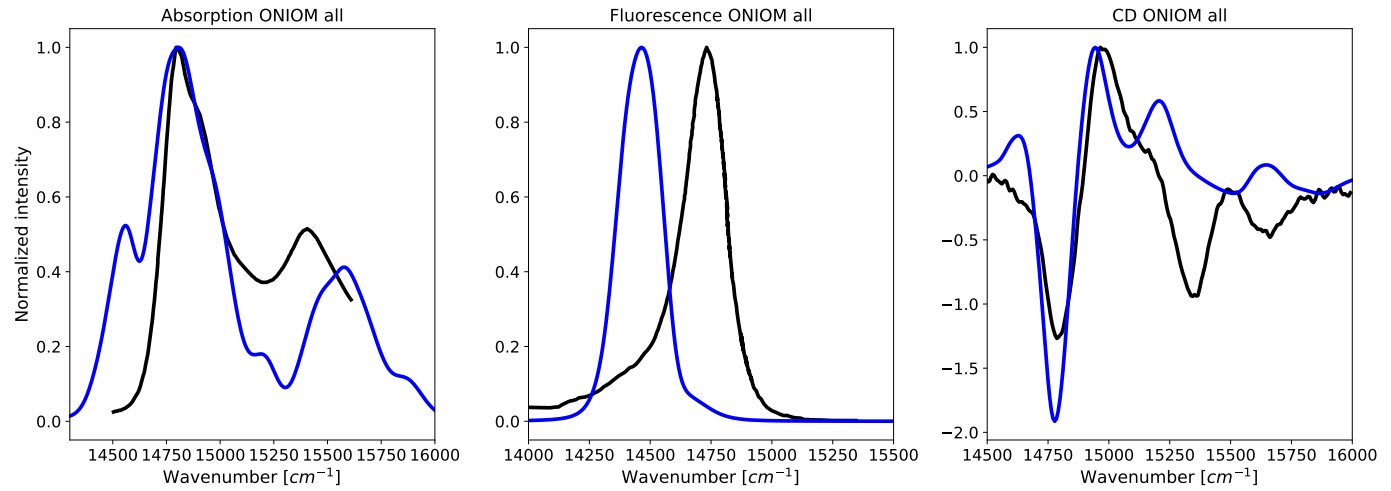


Fig. S4 TDM06-2X/MMPol spectra (at 77K) calculated for the structure where the chlorophylls and axial amino-acids were optimized. The black spectra correspond to the experimental absorption[3], fluorescence and CD spectra[4] at 77K.

Table S3 Intramonomer Hamiltonian calculated within QM/MMpol approach at M062x/6-31+G(d) level. All the values have been obtained as averages on the configurations extracted from MD. The reported couplings are without rescaling. The site excitation energies were shifted by -1210 cm^{-1}. All values are in cm^{-1}.

	b601	a602	a603	a604	b605	b606	b607	b608	b609	a610	a611	a612	a613	a614
b601	15491	87	-8	0	0	0	0	0	0	-6	37	6	-14	3
a602	87	14858	48	9	0	8	8	-7	-31	-17	-5	23	-3	1
a603	-8	48	15004	-2	0	-7	13	6	142	19	-1	-1	3	-9
a604	0	9	-2	15072	2	125	36	-2	-4	-4	-5	4	2	-7
b605	0	0	0	2	15324	31	-6	-7	-1	0	0	0	0	0
b606	0	8	-7	125	31	15261	30	-3	1	-3	0	3	2	0
b607	0	8	13	36	-6	30	15303	-6	-11	1	0	3	4	-4
b608	0	-7	6	-2	-7	-3	-6	15224	39	110	7	1	0	0
b609	0	-31	142	-4	-1	1	-11	39	15418	-5	1	-1	-4	0
a610	-6	-17	19	-4	0	-3	1	110	-5	14857	-42	36	9	0
a611	37	-5	-1	-5	0	0	0	7	1	-42	14998	226	-9	2
a612	6	23	-1	4	0	3	3	1	-1	36	226	15067	3	1
a613	-14	-3	3	2	0	2	4	0	-4	9	-9	3	14974	-69
a614	3	1	-9	-7	0	0	-4	0	0	0	2	1	-69	15075

Table S4 Intermonomer couplings (cm^{-1}) calculated with the QM/MMpol approach at M062x/6-31+G(d) level. The reported couplings are without rescaling. All the values have been obtained as averages on the configurations extracted from MD.

	b601'	a602'	a603'	a604'	b605'	b606'	b607'	b608'	b609'	a610'	a611'	a612'	a613'	a614'
b601	0	-1	-1	3	-1	3	0	10	69	-6	0	0	0	0
a602	0	-1	20	0	0	-1	0	-1	-15	0	0	0	0	0
a603	0	8	-6	0	0	0	3	0	-1	0	0	0	0	0
a604	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b605	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b606	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b607	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b608	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b609	0	0	3	0	0	0	0	0	0	0	0	0	0	0
a610	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a611	0	0	-2	0	0	0	0	3	8	0	0	0	0	0
a612	0	0	0	0	0	0	0	0	-5	0	0	0	0	0
a613	0	0	-10	-3	-1	-2	17	0	6	0	0	0	0	0
a614	0	0	6	-8	5	-10	-4	0	-6	0	0	0	0	0

Table S5 Intramonomer Hamiltonian calculated for the crystal structure within QM/MMpol approach at M062X/6-31+G(d) level. The reported couplings are without dipole rescaling. The site excitation energies were shifted by -1210 cm^{-1}. All values are in cm^{-1}.

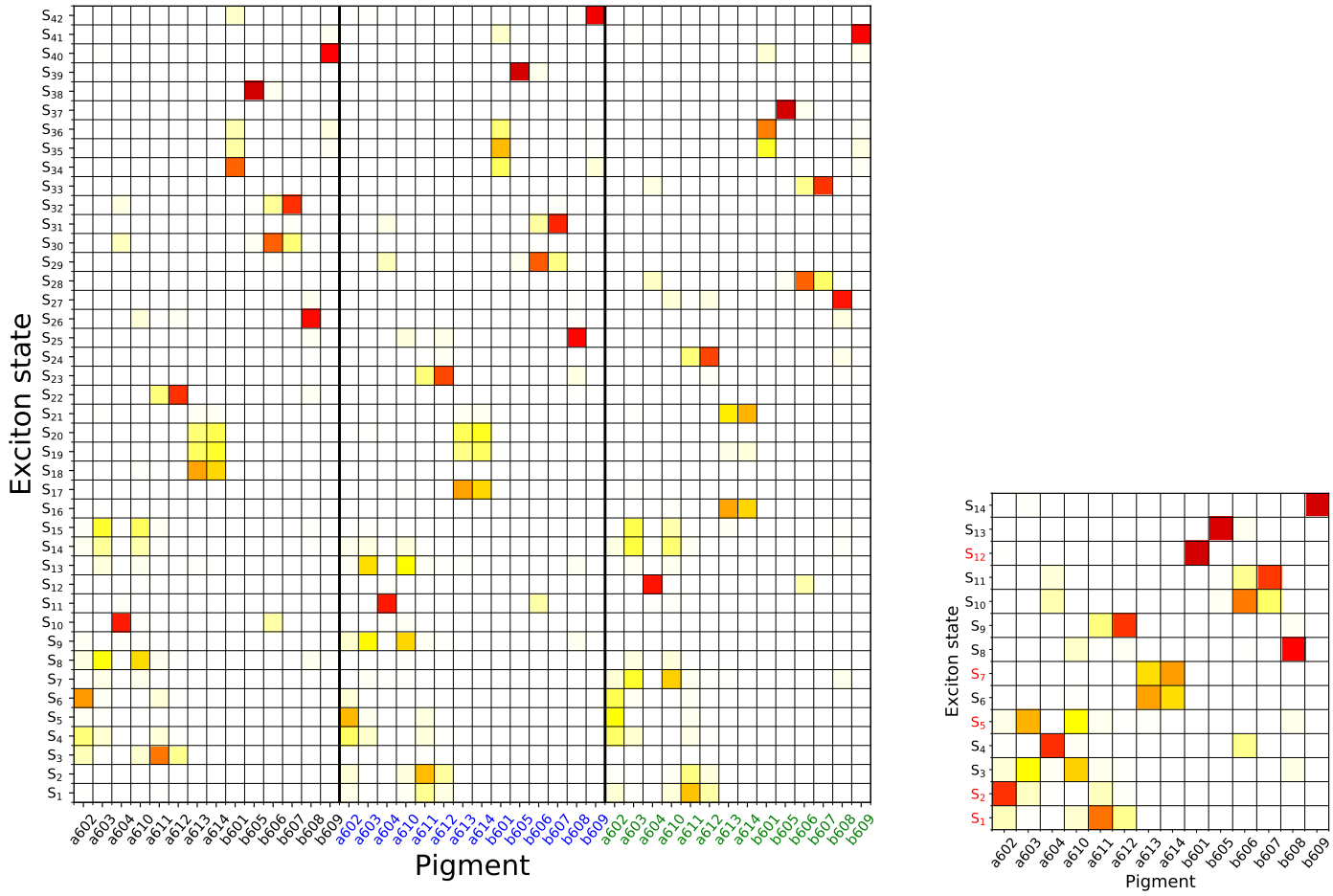
	b601	a602	a603	a604	b605	b606	b607	b608	b609	a610	a611	a612	a613	a614
b601	15584	58	-8	0	0	0	0	0	0	-7	35	5	-16	1
a602	58	14847	33	10	0	9	9	-8	-28	-13	-9	25	0	-1
a603	-8	33	14939	-1	0	-3	13	5	105	19	1	-1	5	-10
a604	0	10	-1	14995	4	159	50	-4	-4	-2	-7	8	7	-7
b605	0	0	0	4	15605	57	-10	-8	-1	0	0	0	0	0
b606	0	9	-3	159	57	15247	47	-9	-3	-1	0	3	3	0
b607	0	9	13	50	-10	47	15384	-4	-6	0	0	3	7	-4
b608	0	-8	5	-4	-8	-9	-4	15176	40	82	9	1	0	0
b609	0	-28	105	-4	-1	-3	-6	40	15773	-6	4	0	-5	0
a610	-7	-13	19	-2	0	-1	0	82	-6	14944	-44	22	12	0
a611	35	-9	1	-7	0	0	0	9	4	-44	14876	210	-4	1
a612	5	25	-1	8	0	3	3	1	0	22	210	15141	0	1
a613	-16	0	5	7	0	3	7	0	-5	12	-4	0	15037	-56
a614	1	-1	-10	-7	0	0	-4	0	0	0	1	1	-56	15049

Table S6 Average intermonomer couplings calculated for the crystal structure within QM/MMpol approach at M062X/6-31+G(d) level. The reported couplings are without rescaling. All values are in cm^{-1}.

	b601'	a602'	a603'	a604'	b605'	b606'	b607'	b608'	b609'	a610'	a611'	a612'	a613'	a614'
b601	0	0	2	4	-3	3	3	14	63	-7	0	0	0	0
a602	0	-1	18	0	0	-1	-1	-1	-14	0	0	0	0	0
a603	0	9	-11	0	0	0	2	0	0	0	0	0	0	0
a604	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b605	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b606	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b607	0	0	-3	0	0	0	0	0	0	0	0	0	0	0
b608	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b609	0	0	5	0	0	0	0	0	0	0	0	0	0	0
a610	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a611	0	0	-2	0	0	-1	0	5	9	0	0	0	0	0
a612	0	0	0	0	0	0	0	0	-5	0	0	0	0	0
a613	0	0	-11	1	-2	0	23	0	6	0	0	0	0	0
a614	0	0	7	-8	7	-10	-5	0	-6	0	0	0	0	0

Table S7 TDM06-2X/MMPol exciton energies (in cm^{-1}) determined from MD structures and comparison with the same energies calculated by Müh et al.[2], Novoderezhkin et al.[1] and determined from 2D optical spectroscopy by Calhoun et al.[7] and by Do et al.[8]. The exciton energies from the current work were determined from the averaged LHCII monomeric Hamiltonian in Table S3.

	Current work	Müh et al.	Novoderezhkin et al.	Calhoun et al.	Do et al.
1	14748	14750	14699	14700	14685-14705
2	14833	14840	14751	14770	14750-14770
3	14867	14860	14804	14810	14815
4	14940	14860	14858	14880	14860
5	14983	14880	14918	14910	14900-14925
6	15008	14900	14952	14990	14970-15015
7	15110	14990	14992	15030	15060-15105
8	15245	15040	15022	15130	
9	15260	15175	15210	15210	15265
10	15262	15290	15306	15290	
11	15325	15400	15363	15360	
12	15367	15400	15416	15430	
13	15470	15555	15456	15480	
14	15509	15650	15512	15510	



(a) **(b)**

Fig. S5 Comparison of the the exciton composition calculated at the TDM06-2X/MMPol for the crystal structure of the a) LHCII trimer and b) the single monomer (presented in the main text). Each square represents the contribution of a pigment to an exciton state for a single monomer (obtained through the probability coefficient $|c_i^{(M)}|^2$). Red squares mean high contribution, white ones negligible contribution. The numbering of the exciton states refers to a trimer (42 states) and a monomer (14 states), respectively. The labels reported in red for the monomer indicate the states with more than 25% contribution of delocalization over more than one monomer (see trimeric exciton composition). For the trimer a different color is used for the pigment labels of each monomer. In the LHCII trimer the delocalization of the exciton states over more monomers is present, but monomeric contributions are clearly recognizable.

S4 Effect of the dipole rescaling

The transition dipole strength calculated for Chl-b in vacuo is in agreement with the experimental estimate.[9] The agreement between calculations and experiments is worse for Chl-a. In order to investigate the effect of this inaccuracy in our calculations, the size of the Chl-a transition dipole was set to $\alpha\mu_{QM}^{chl-a}$, where $\alpha = \left(\mu_{QM-VAC}^{chl-b}/\mu_{QM-VAC}^{chl-a}\right)\cdot\sqrt{14.7/21.0}$: introducing such a scaling factor we in fact obtain the experimental ratio between Chl-a and Chl-b transition dipole strengths, where $14.7D^2$ correspond to 0-0 dipole strength of Chl-a and $21.0D^2$ the one for Chl-b. The comparison of the optical spectra calculated with and without rescaling is presented in the Fig. S6. As it can be seen the effects are quite small.

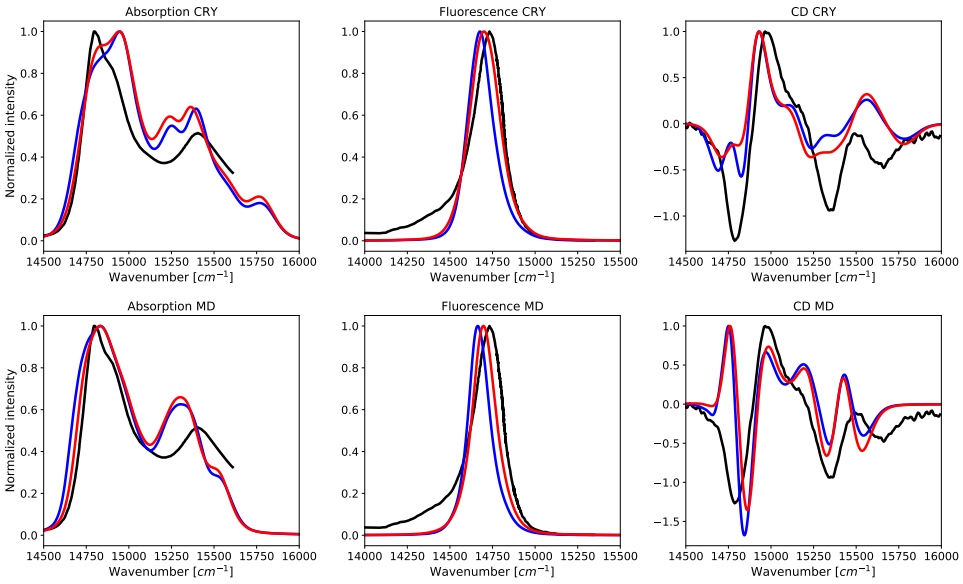


Fig. S6 Optical spectra at 77K calculated with spectral density from Novoderezhkin et al. The blue line represent the optical spectra calculated without the dipole rescaling and the red one with the dipole rescaling. The black spectra correspond to the experimental absorption [3], fluorescence and CD spectra [4].

S5 Environmental contribution to site energies

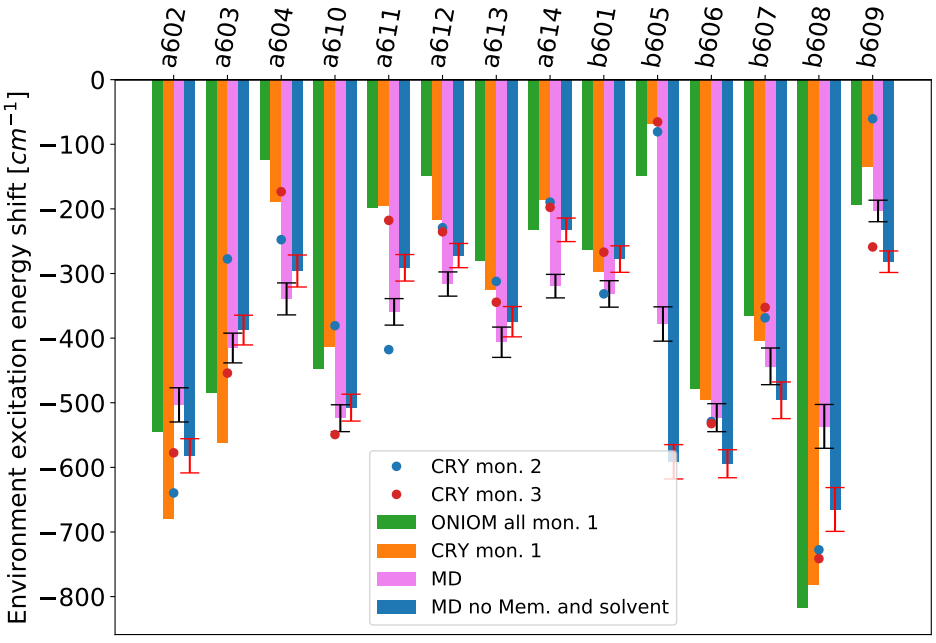


Fig. S7 Comparison of the environment induced shifts in the excitation energies for the crystal structure (CRY), the optimized Chl geometries within the protein in the crystal structure (ONIOM all) and the averages from the MD configurations. The black and red error bars represent the 95% confidence intervals for the mean value of the environmental shift.

Table S8 Table of the z-scores ($z = (\Delta E_{CRY} - \Delta E_{MD})/\sigma_{\mathrm{env}}$) from the MD distribution of the environmental excitation energy shifts for the crystal and optimized LHCII structures. See Table S9 for the values of σ_{env}

	CRY mon. 1	CRY mon. 2	CRY mon. 3	ONIOM all
a602	-1.078	-0.832	-0.453	-0.256
a603	-1.031	0.968	-0.272	-0.488
a604	0.983	0.599	1.083	1.404
a610	0.855	1.112	-0.197	0.598
a611	1.285	-0.46	1.114	1.266
a612	0.856	0.752	0.696	1.44
a613	0.56	0.649	0.427	0.866
a614	1.183	1.154	1.082	0.767
b601	0.274	0.001	0.51	0.535
b605	1.887	1.809	1.902	1.397
b606	0.208	-0.045	-0.069	0.334
b607	0.225	0.427	0.519	0.441
b608	-1.171	-0.909	-0.975	-1.34
b609	0.657	1.384	-0.54	0.086

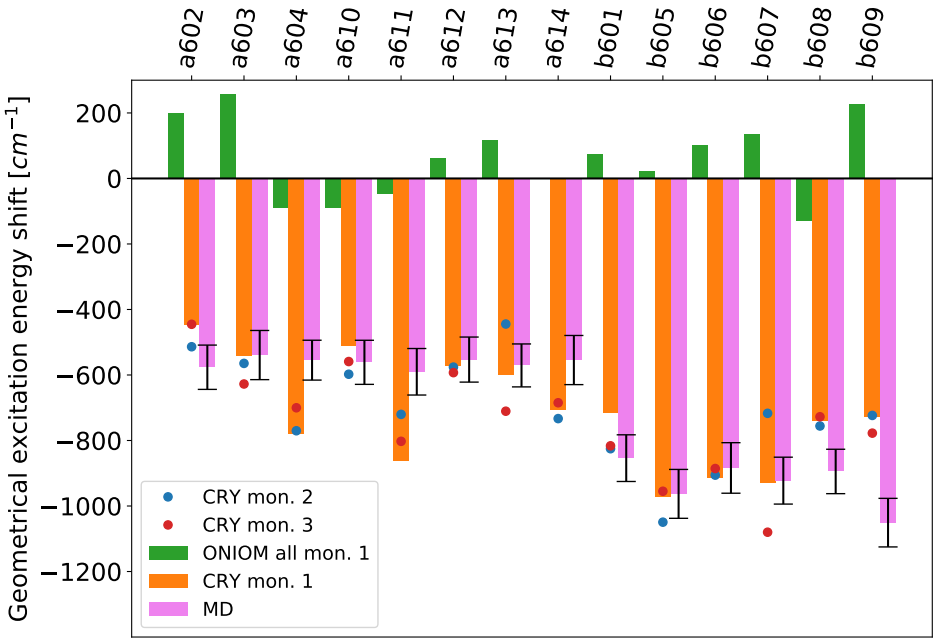


Fig. S8 Comparison of the geometrical contributions to the excitation energy shifts for the crystal structure (CRY), the optimized Chl geometries within the protein in the crystal structure (ONIOM all) and the averages from the MD configurations. The black error bars represent the 95% confidence intervals for the mean value of the geometrical energy shift.

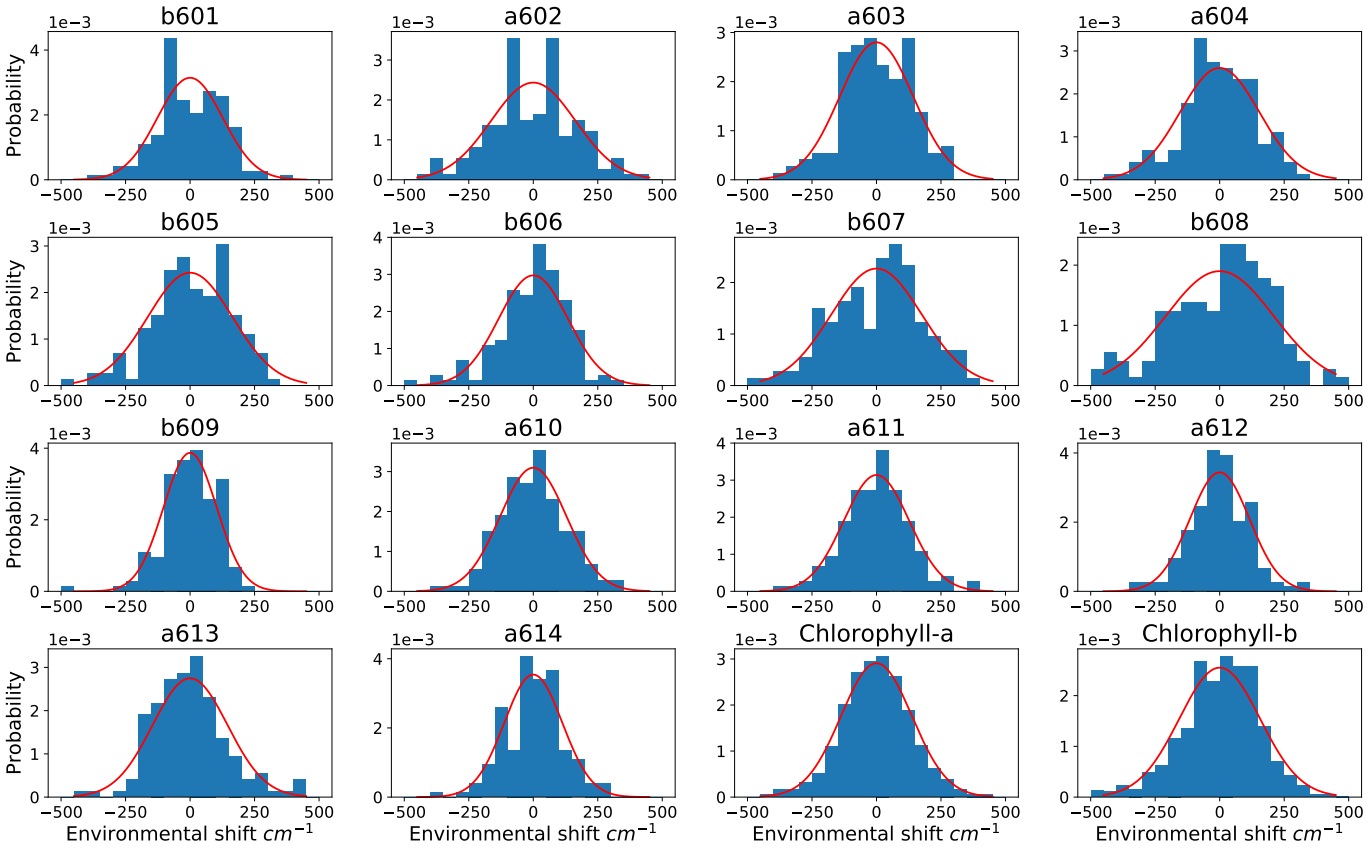


Fig. S9 Distribution of the environment-induced shifts in the site energies (in cm^{-1}) from the MD simulation at 300K. The distribution was obtained from the 150 realizations (50 conformations of the trimer). The standard deviations are presented in the Table S9 in the main text

Table S9 Environment disorder (σ_{env}) obtained from the distributions of site energy shifts in Fig. S9

$T = 300K$	$\sigma\ [cm^{-1}]$		$\sigma\ [cm^{-1}]$		$\sigma\ [cm^{-1}]$
b601	127	b606	134	a611	127
a602	164	b607	176	a612	116
a603	142	b608	210	a613	145
a604	153	b609	103	a614	113
b605	164	a610	129		

References

1 V. I. Novoderezhkin, M. A. Palacios, H. van Amerongen and R. van Grondelle, *The Journal of Physical Chemistry B*, 2004, **108**, 10363–10375.

2 F. Müh, M. E.-A. Madjet and T. Renger, *The Journal of Physical Chemistry B*, 2010, **114**, 13517–13535.

3 F. J. Kleima, C. C. Gradinaru, F. Calkoen, I. H. M. van Stokkum, R. van Grondelle and H. van Amerongen, *Biochemistry*, 1997, **36 49**, 15262–8.

4 P. W. Hemelrijk, S. L. Kwa, R. van Grondelle and J. P. Dekker, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1992, **1098**, 159–166.

5 S. Caffarri, R. Croce, J. Breton and R. Bassi, *The Journal of biological chemistry*, 2001, **276**, 35924–33.

6 P. H. Lambrev, Z. Várkonyi, S. Krumova, L. Kovács, Y. Miloslavina, A. R. Holzwarth and G. Garab, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 2007, **1767**, 847 – 853.

7 T. R. Calhoun, N. S. Ginsberg, G. S. Schlau-Cohen, Y.-C. Cheng, M. Ballottari, R. Bassi and G. R. Fleming, *The Journal of Physical Chemistry B*, 2009, **113**, 16291–16295.

8 T. N. Do, A. Huerta-Viga, P. Akhtar, H. L. Nguyen, P. J. Nowakowski, M. F. Khyasudeen, P. H. Lambrev and H.-S. Tan, *The Journal of Chemical Physics*, 2019, **151**, 205101.

9 R. S. Knox and B. Q. Spring, *Photochem. Photobiol.*, 2003, **77**, 497–501.