Absorption spectra of FAD embedded in cryptochromes (Supporting Information)

Claus Nielsen,* Morten S. Nørby, Jacob Kongsted, and Ilia A. Solov'yov*

Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense M, Denmark

E-mail: clausnielsen@sdu.dk; ilia@sdu.dk

Additional method details

Equilibration of cryptochrome structures

The structure of a protein fluctuates due to the thermal motion of the atoms in the protein itself and its immediate environment. These fluctuations affect the absorption spectrum, and, therefore, multiple structures of each cryptochrome type were considered. The molecular structures for various cryptochromes used for the absorption spectra calculations were obtained from earlier molecular dynamics (MD) simulations (1–5) obtained by us, see Table 1 for a summary. The production simulations for each cryptochrome structure delivered 10 snapshots sampled at 10 ns intervals, which were then used as the structures for spectra calculations. The sampling interval of 10 ns was used to ensure that the 10 structures were statistically independent, and a total of 100 ns of the production simulations was, therefore, used to generate structures for the spectra calculations. As dicussed below, 10 structures per cryptochrome seemed sufficient for the present study.

Molecular dynamics simulations of MmCry

The crystal structure of MmCry (6) does not contain an FAD co-factor, and as such the co-factor was put there *in silico*. The backbone of MmCry was aligned with the backbone of an FAD-containing AtCry1 structure, and the FAD molecule from the AtCry1 structure was then added into MmCry. Extensive energy minimization and equilibration simulation was then performed on the FAD-containing MmCry structure, following essentially the same protocol as ErCry4 (4).

Simulation details

The MD trajectories employed in the present investigation were taken from several previous studies (1–5), and therefore do not follow exactly the same simulation protocol. The most

essential details of the simulations are summarized in Table S1 below, and additional information can be found in the original papers (1–5). The simulations used a 2 fs integration time step with a Langevin thermostat, as well as a Langevin barostat with a pressure control set to 1 atm for the NPT simulations. All proteins were solvated in a water box with periodic boundary conditions. The van der Waals interactions had a smooth cutoff distance of 12 Å, and the electrostatic interactions were evaluated using the particle mesh Ewald (PME) approach in all simulations (7). All simulations except for the one with XlCry were perfomed using NAMD (8) with CHARMM forcefields including CMAP corrections (2, 9–11), while the XlCry simulation was carried out using GROMACS (12) with the Amber99SB (13, 14) force field for the protein and generalized Amber force field (GAFF) for FAD (2, 14). The ShakeH algorithm (15) was used in NAMD to constrain bond lengths for hydrogen atoms, while an analogous LINCS algorithm (16) was used in the GROMACS simulation.

Table S1: **Summary of the MD simulations protocols.** The essentials of the earlier MD simulations (1–5) of the six cryptochromes. The simulations were carried out completely independent from each other, and, therefore, follow slightly different simulation protocols; see original papers for more details.

System	DmCry	AtCry1	ErCry1	ErCry4	MmCry	XlCry
Simulation length (ns)	135	107	310	1012	540	135
Ensemble	NVT	NVT	NPT	NPT	NPT	NPT
Temperature (K)	300	300	300	310	300	310
Crystal structure ref.	(17, 18)	(19)	-	-	(6)	-
PDB ID	4GU5	1U3C	-	-	4K0R	-
Simulation ref.	(1)	(2, 3)	(1)	(4)	this study	(5)

Structure optimization

The structures taken directly from MD simulations typically reproduce bond lengths, valence angles and dihedral angles in the structure only qualitatively. These parameters, however, are crucial for determining an absorption spectrum, which calls for geometry optimization of the core quantum region. In all the studied cryptochromes, the core region was defined

as the isoalloxazine part of the FAD co-factor, as illustrated in Fig. 1. The QM/MM geometry optimization was carried out using the QSite module of the Maestro application of the Schrödinger suite (20), where a hydrogen cap was used to terminate the dangling bonds on the interface between the quantum region and its environment. All atoms outside the quantum region were kept frozen during this optimization. The QM part of the optimization was performed using density functional theory (DFT) with the B3LYP functional (21, 22) and the lacvp* basis set (which is identical to 6-31G* for all the studied atoms) (23), while the OPLS 2005 force-field (24) was employed for the MM part.

Comparison with experimental spectra

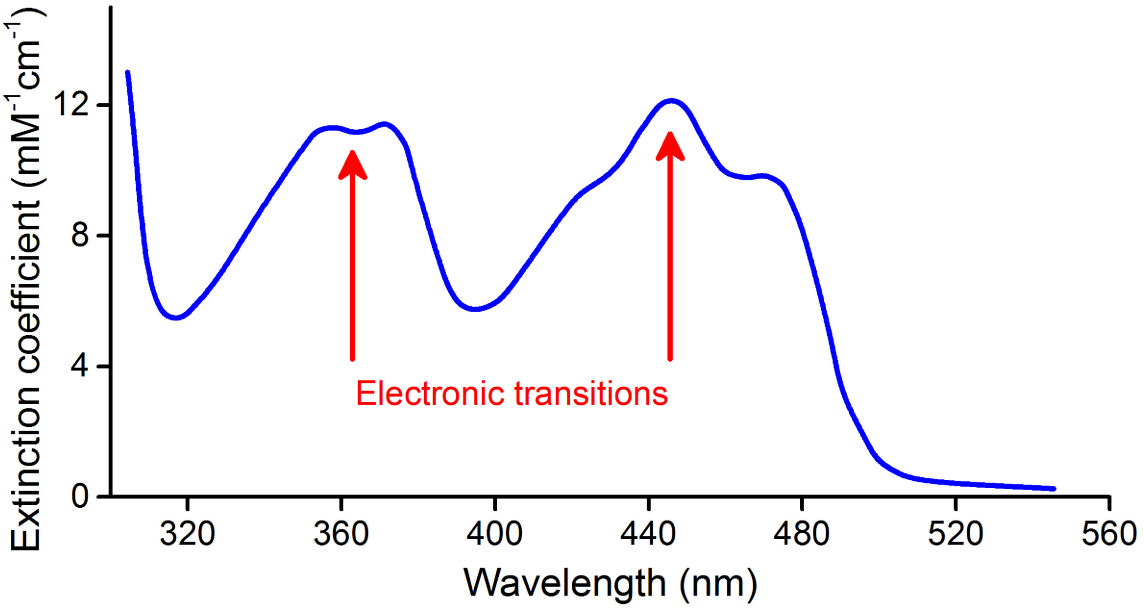

Figure S1: **Experimental absorption spectrum for FAD in AtCry1.** The experimentally obtained absorption spectrum has overall similar spectral features as the calculated spectrum. The electronic transitions are marked with red arrows; those are the transitions that were calculated and therefore correspond to the peaks in Fig. 2. The experimental data was adapted from (25).

The experimental spectrum of AtCry1 has a prominent electronic transition at 450 nm,

see Fig. S1, and the calculated AtCry1 spectrum had a corresponding peak at 390 nm (see results). The difference in the peak position of the calculated spectrum relative to experiment is within the expected error range of a DFT calculation (26). Thus the calculated AtCry1 spectrum is shifted by 60 nm towards longer wavelengths to allow direct comparison of the calculated and experimental spectra, and since the quantum region is identical for all the studied cryptochromes, a shift of 60 nm was, therefore, applied to all the obtained spectra.

Supplementary results

Stability of the studied structures

The stability of the equilibrated structures can be verified by considering the root mean square deviation (RMSD) of the studied structure (or its components). Such an analysis has been performed for the protein backbone of the six studied cryptochromes and for the FAD cofactor. The time evolution of the RMSDs in both cases is shown in Figs. S2 and S3, respectively.

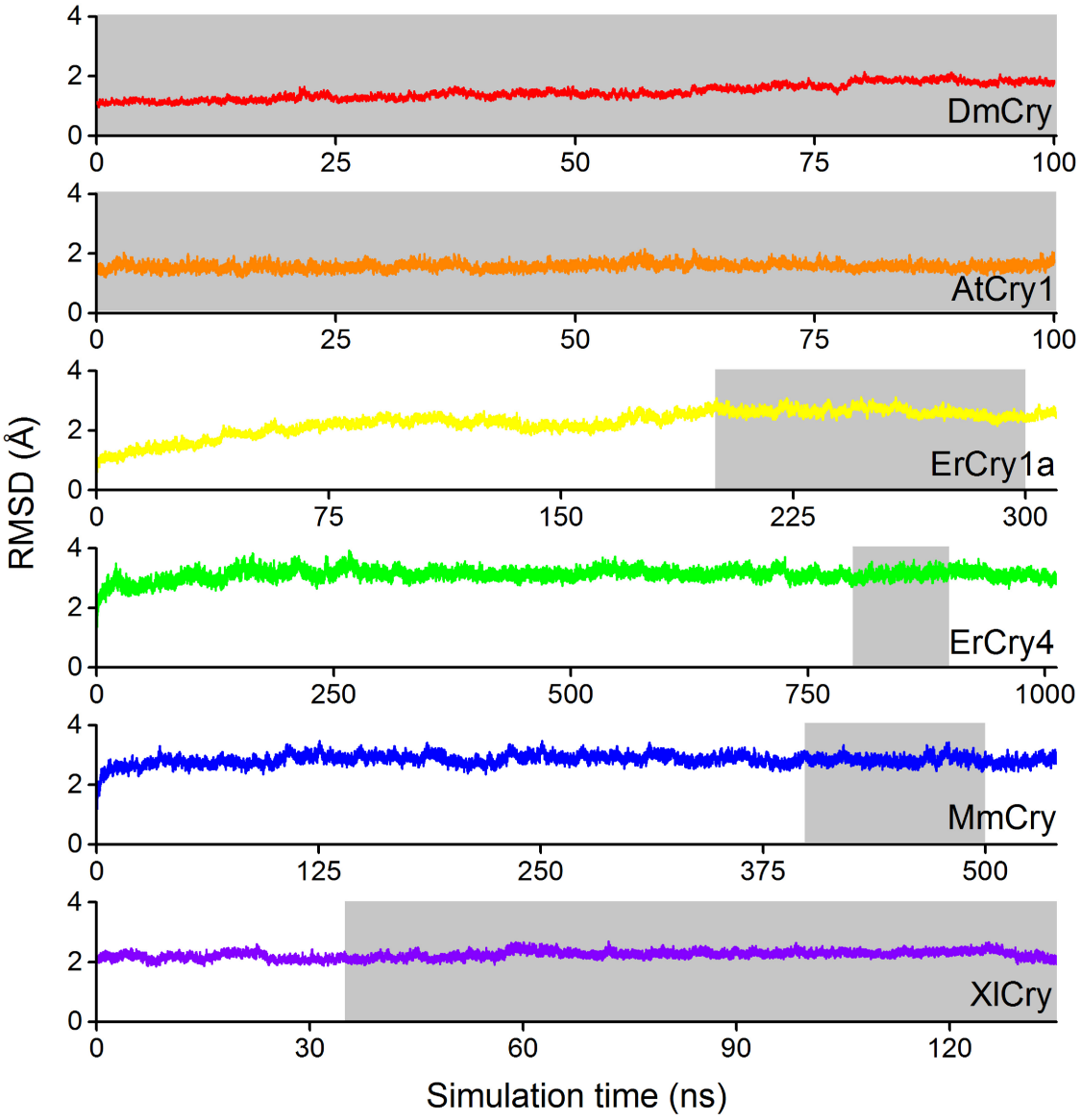

Figure S2: **RMSD of the protein backbone.** For each studied cryptochrome, the structure of the protein backbone at every frame of the simulation was aligned with the corresponding initial structure at the beginning of the production simulation. The snapshots used to calculate the absorption spectra of FAD are taken from the shaded regions of 100 ns duration; 10 snapshots are taken at a 10 ns interval.

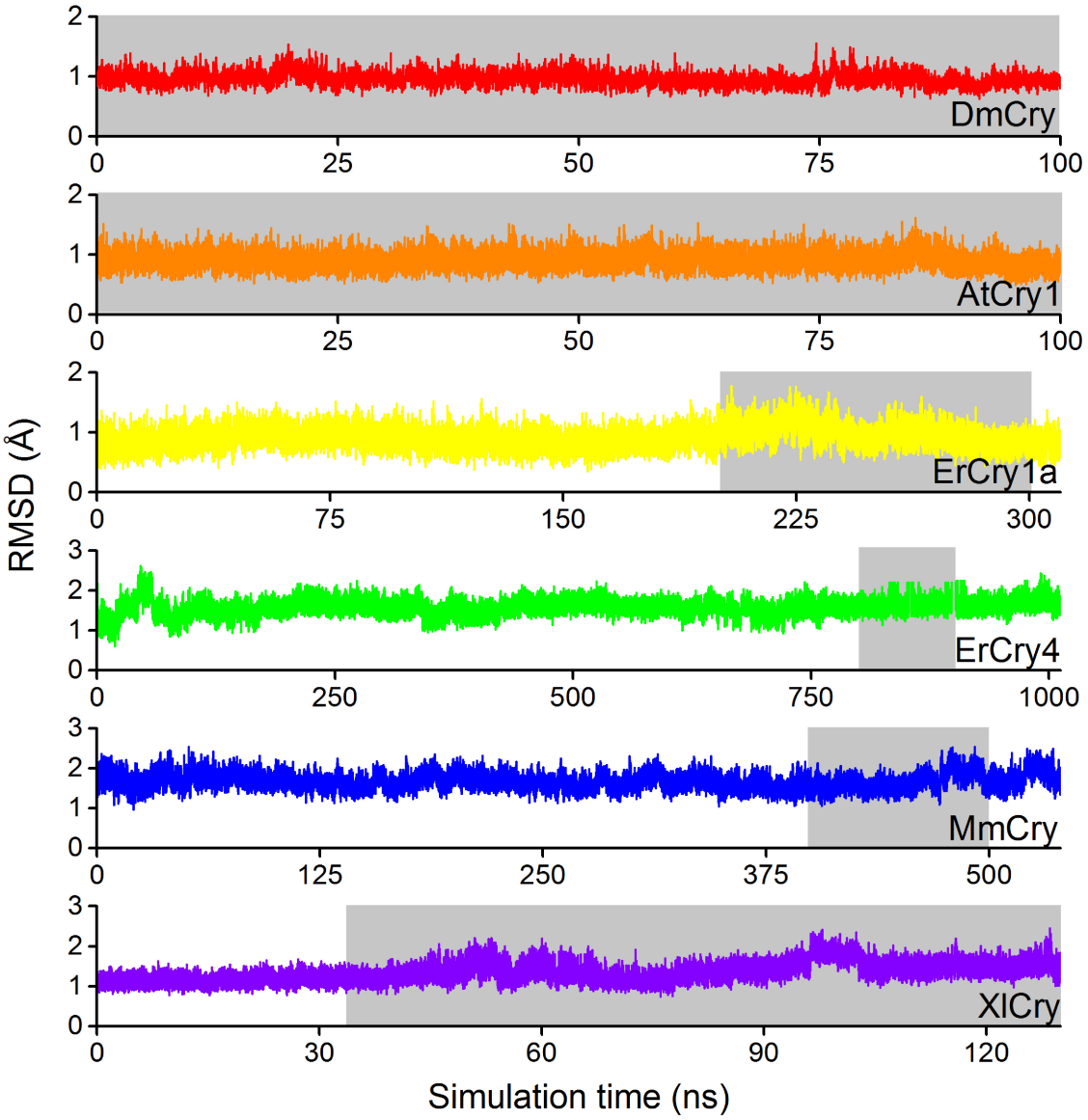

Figure S3: **RMSD of the FAD cofactor.** For each studied cryptochrome, the structure of the protein backbone at every frame of the simulation was aligned with the corresponding initial structure at the beginning of the production simulation. The snapshots used to calculate the absorption spectra of FAD are taken from the shaded regions of 100 ns duration; 10 snapshots are taken at a 10 ns interval.

Full FAD in the quantum (QM) region

For all the spectra calculations only the isoalloxazine moiety (30 atoms of FAD) was treated quantum mechanically, i.e. as part of the quantum (QM) region, while the rest of FAD together with the surrounding protein and some water and ions were considered the environment and described in terms of a polarizable embedding potential. This choice of the QM region is fine insofar as the relevant electronic transitions happen between electronic states that are completely localized on the isoalloxazine moiety, but the calculated spectra would be wrong if this assumption does not hold. Thus in order to justify the choice of QM region, a calculation that includes the entire FAD in the QM region was performed. The result of this calculation with FAD in the QM region and everything else still represented by a polarizable embedding potential is shown in Fig. S4; the calculation is based on a single structure of AtCry1, and the corresponding spectrum with only isoalloxazine in the

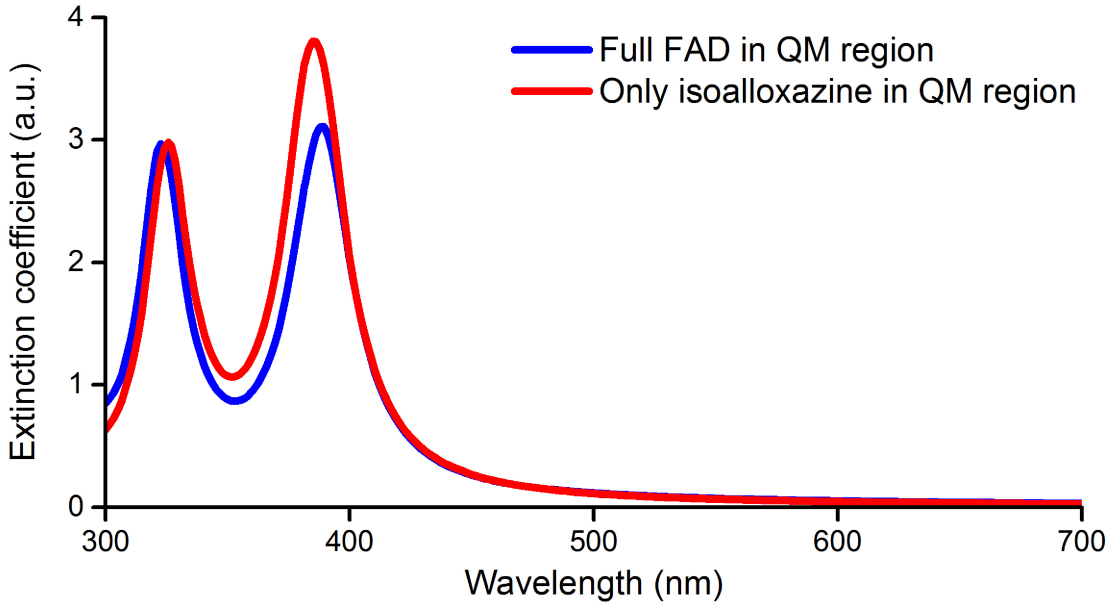

Figure S4: **Spectra of AtCry1 with different QM regions.** Blue: Entire FAD is part of the QM region. Red: Only the 30 atoms of the isoalloxazine moiety is included in the QM region, while the rest of FAD is included in the polarizable embedding (PE) potential instead. The protein environment is described by a PE potential in both calculations, which were performed on a single AtCry1 structure. The wavelengths have not been shifted.

QM region for the same structure is shown for comparison.

There is a small change in intensity of the primary peak (∼400 nm) in Fig. S4, but the peak positions remain almost unchanged. Thus including the rest of FAD in the QM region does not appear to have any significant effect on the calculated absorption spectrum.

Convergence of the absorption spectra

The performed spectra analysis is based on 10 stastistically independent structures sampled at 10 ns intervals from MD trajectories, see Table. S1. The same procedure was used for all the six cryptochromes investigated here. In order to show that 10 independent structures are sufficient to obtain a statistically converged spectrum of the isoalloxazine moiety in different cryptochromes, 10 additional spectra calculations for ErCry4 were performed. The results

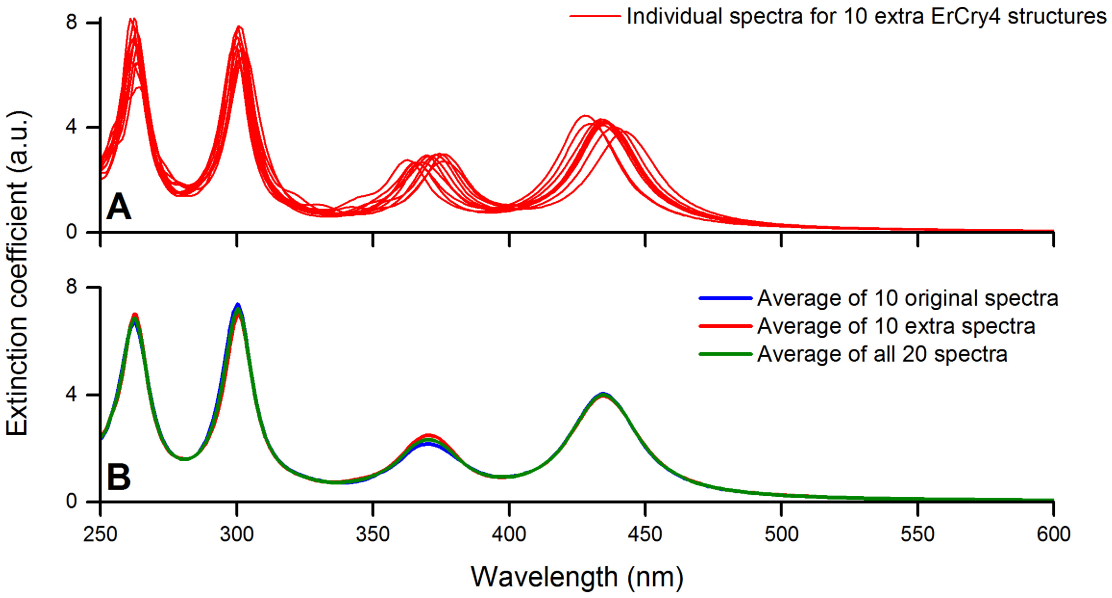

Figure S5: **Convergence of ErCry4 spectra calculations.** **A**: Individual spectra obtained for 10 ErCry4 structures in addition to the original 10 ones, see Fig. 2. **B**: The average spectra computed for the original 10 ErCry4 structures (blue), the 10 additional structures (red) and the average spectrum of all 20 structures (green); inclusion of the 10 additional structures does not change the spectrum significantly. All the spectra have been shifted by a systematic error correction of 60 nm, as also done in the plots shown in Figs. 2 and 3.

in Fig. S5 shows that while there is some variance among the individual spectra of the 10 additional structures, the average spectrum of all 20 structures is not changed significantly by considering the 10 additional structures.

Environment of the isoalloxazine moiety of FAD

The difference between the calculated absorption spectra of the six different cryptochromes is caused by the difference in the electrostatic potentials produced by the protein matrices. Figure S6 visualizes this difference, and features the electrostatic potentials of the protein matrices projected onto a solid-surface representation of the isoalloxazine moiety of FAD. The electrostatic potential was calculated using the PMEPot plugin in VMD (27), where the potential was averaged over the 10 structures per cryptochrome, used in the spectra calculations. Thus the classic force fields used in the MD simulations were used to calculate the electrostatic potentials.

In order to understand the differences in electrostatic potentials between the cryptochromes, one should look at the amino acid residues near the isoalloxazine moiety. Figure S7 shows the sidechains of all the amino acids that have at least a single atom within a distance of 5 Å away from the isoalloxazine moiety, for each of the six different cryptochromes. The residues have been color coded, such that blue indicates positively charged, red indicates negatively charged, green indicates neutral but polar, and white indicates neutral and unpolar residues. The only exception is D396(H) in AtCry which is drawn in red despite being protonated and, therefore, effectively neutral.

FAD in ErCry4 and XlCry, and to a less extend DmCry, appears to feel a negative electrostatic potential from the protein matrix as evidenced in Fig. S6, whereas the other three cryptochromes feels a more positive electrostatic environment around FAD. Interestingly, as follows from Fig. S7, there does not appear to be a significantly different number of charged amino acids in the six studied cryptochrome structures, however, the structures showing a

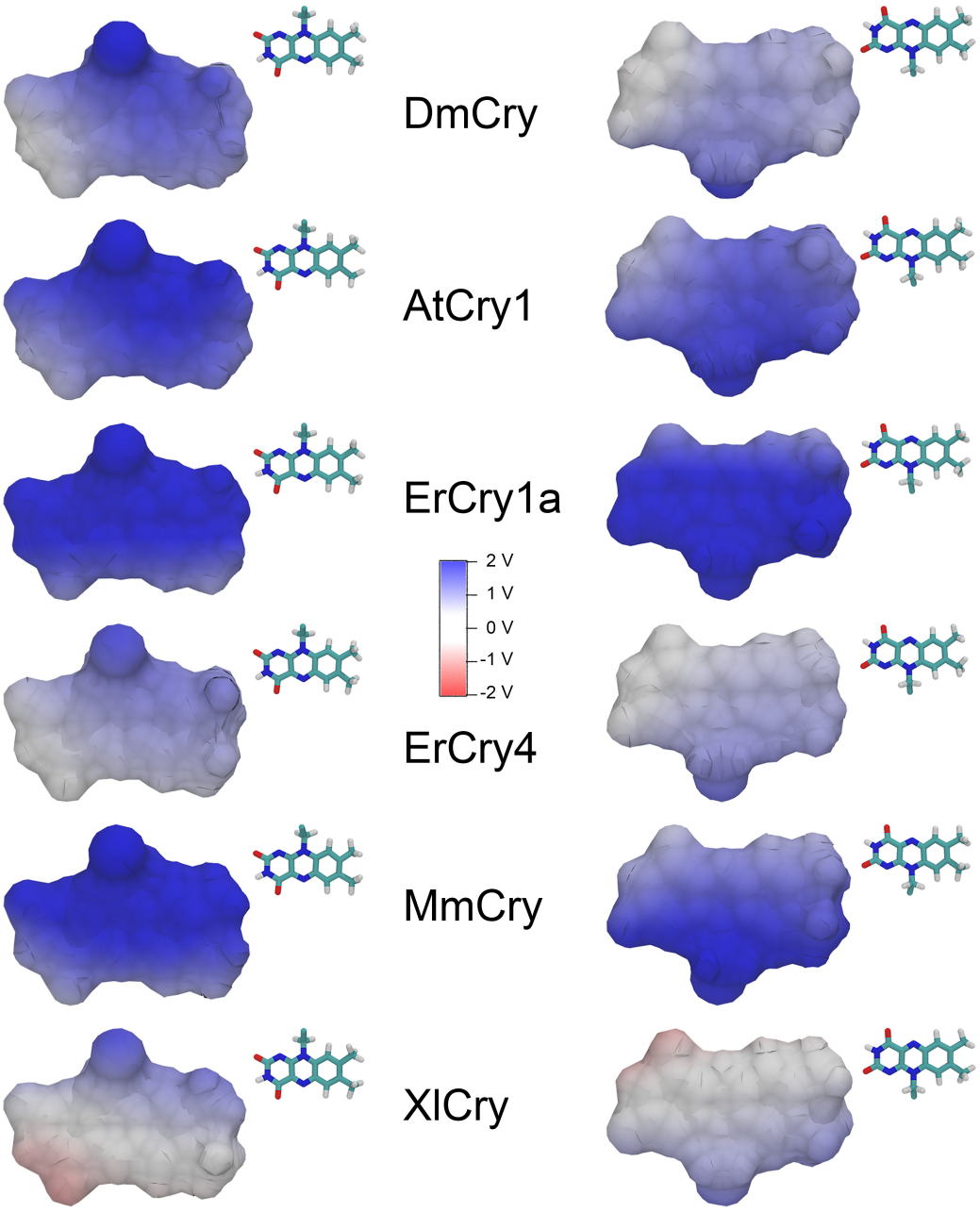

Figure S6: **Local electrostatic environment of the isoalloxazine moiety of FAD.** Electrostatic potential of the protein matrix averaged over the 10 structures used for the spectrum calculations. The electrostatic potential is projected onto a surface containing the isoalloxazine moiety of FAD, and the potential is shown from both of its sides. The small atomistic representations of the FAD isoalloxazine moiety shows how the moiety is oriented inside the surface. The potential is calculated from the forcefield parameters used in the MD simulations. Note that no part of FAD itself contributes to the calculated potentials.

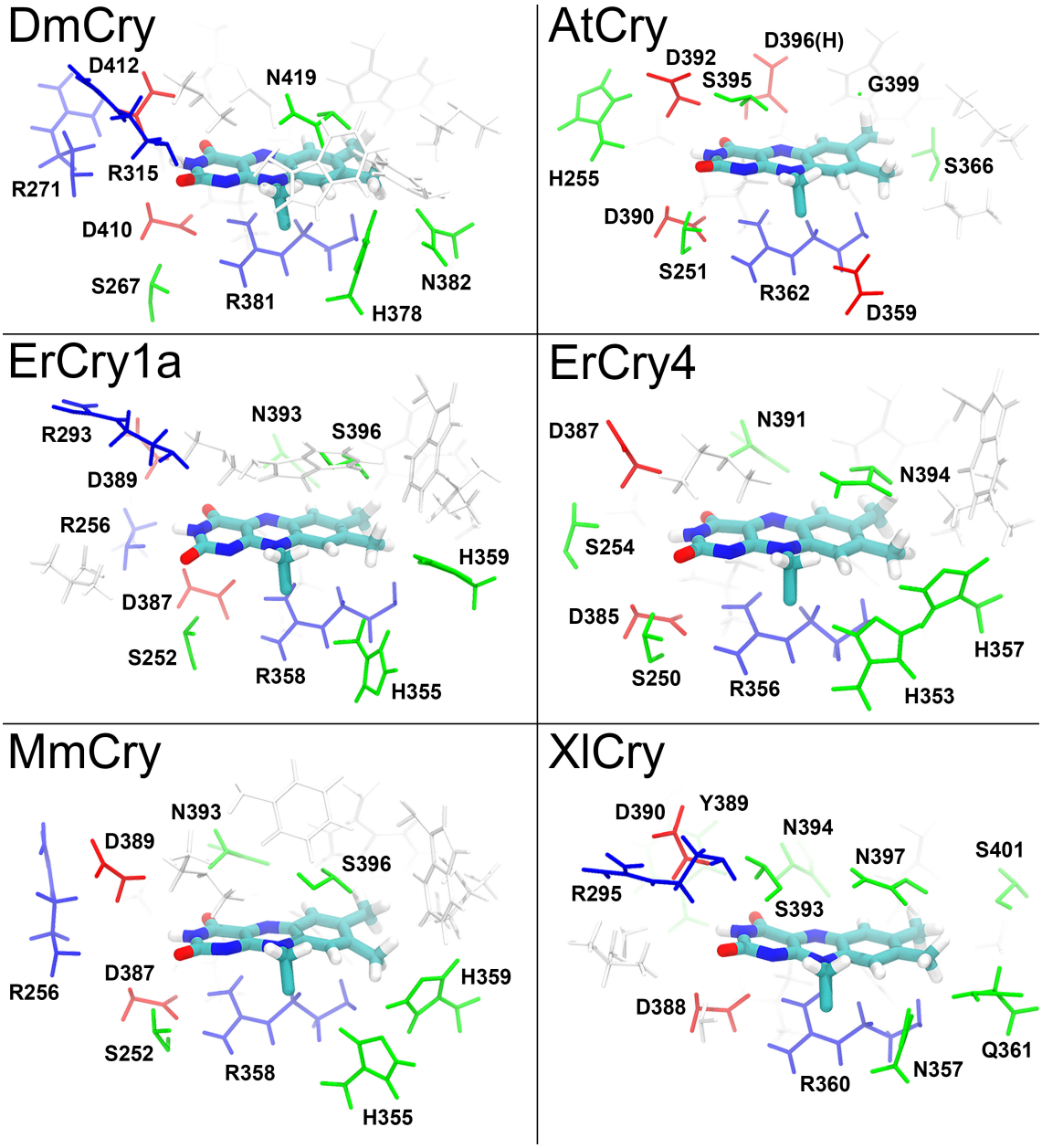

Figure S7: **Amino acid residues located within 5 Å of the isoalloxazine moiety of FAD.** Molecular rendering of the FAD binding pocket in the six studied cryptochrome structures, corresponding to the equilibrated configurations of the protein, see Fig. S2 and S3. Only the sidechains of the residues are shown, and are colored by type: charged positive (blue), charged negative (red), neutral polar (green) and neutral non-polar (white).

negative electrostatic potential around the FAD binding cavity appear to have a larger number of uncharged polar amino acids located near the isoalloxazine moiety. The differences in

the nearby residues are too complex to provide a simple explanation for the differences in electrostatic potential among the six cryptochromes; the only consistent feature is a positive potential near the carbon atom that connects the isoalloxazine moiety with the remainder of FAD, which seems to be caused by a nearby arginine that is present in all six cryptochromes, i.e. R358 in ErCry1a and MmCry, R381 in DmCry, etc.

Spectrum of FADH

A spectrum of FADH embedded in AtCry1 in Fig. 4 of the main paper was calculated following the same protocol as the spectra of FAD, except that only a single structure was used; the calculated FADH spectra was not averaged over 10 structures as it is only meant as a reference calculation. The studied structure was taken from an equilibrated MD trajectory of AtCry1 with FADH (28). Also note that the quantum region for the FADH spectrum calculation was different as it had an additional proton attached to the N5 nitrogen atom, see Fig. 1, and the systematic error in the DFT calculated was, therefore, also different, leading to a shift of 9.5 nm rather than 60 nm towards higher wavelengths.

References

(1) Nielsen, C.; Kattnig, D. R.; Sjulstok, E.; Hore, P. J.; Solov'yov, I. A. Ascorbic acid may not be involved in cryptochrome-based magnetoreception. *J. R. Soc. Interface* **2017**, *14*, 20170657.

(2) Solov'yov, I. A.; Domratcheva, T.; Moughal Shahi, A. R.; Schulten, K. Decrypting Cryptochrome: Revealing the Molecular Identity of the Photoactivation Reaction. *J. Am. Chem. Soc.* **2012**, *134*, 18046–18052.

(3) Solov'yov, I. A.; Domratcheva, T.; Schulten, K. Separation of photo-induced radical pair in cryptochrome to a functionally critical distance. *Sci. Rep.* **2014**, *4*, 3845.

(4) Günther, A.; Einwich, A.; Sjulstok, E.; Feederle, R.; Bolte, P.; Koch, K.-W.; Solov'yov, I. A.; Mouritsen, H. Double-Cone Localization and Seasonal Expression Pattern Suggest a Role in Magnetoreception for European Robin Cryptochrome 4. *Curr. Biol.* **2018**, *28*, 211–223.

(5) Sjulstok, E.; Lüdemann, G.; Kubař, T.; Elstner, M.; Solov'yov, I. A. Molecular insights into variable electron transfer in amphibian cryptochrome. *Biophys. J.* **2018**, *114*, 2563–2572.

(6) Czarna, A.; Berndt, A.; Singh, H. R.; Grudziecki, A.; Ladurner, A. G.; Timinszky, G.; Kramer, A.; ; Wolf, E. Structures of Drosophila Cryptochrome and Mouse Cryptochrome1 Provide Insight into Circadian Function. *Cell* **2013**, *153*, 1394–1405.

(7) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089.

(8) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comp. Chem.* **2005**, *26*, 1781–1802.

(9) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi1 and chi2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(10) MacKerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproduc-

ing protein conformational distributions in molecular dynamics simulations. *J. Comp. Chem.* **2004**, *25*, 1400–1415.

(11) MacKerell, A. D., Jr. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(12) Spoel, D. V. D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.; Berendsen, H. GROMACS: fast, flexible, and free. *J. Comp. Chem.* **2005**, *26*, 1701–1718.

(13) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.

(14) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(15) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(16) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(17) Zoltowski, B.; Vaidya, A.; Top, D.; Widom, J.; Young, M.; Crane, B. Structure of full-length Drosophila cryptochrome. *Nature* **2011**, *480*, 396–399.

(18) Zoltowski, B.; Vaidya, A.; Top, D.; Widom, J.; Young, M.; Crane, B. Corrigendum: Structure of full-length Drosophila cryptochrome. *Nature* **2013**, *496*, 252.

(19) Brautigam, C. A.; Smith, B. S.; Ma, Z.; Palnitkar, M.; Tomchick, D. R.; Machius, M.; Deisenhofer, J. Structure of the photolyase-like domain of cryptochrome 1 from *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12142–12147.

(20) QSite, Schrödinger, LLC, New York, NY, 2017.

(21) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(22) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density. *Phys. Rev.* **1988**, *37*, 785–789.

(23) Hay, P. J.; Wadt, W. R. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* **1985**, *82*, 299–310.

(24) Banks, J. L. et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comp. Chem.* **2005**, *26*, 1752–1780.

(25) Müller, P.; Bouly, J.-P.; Hitomi, K.; Balland, V.; Getzoff, E. D.; Ritz, T.; Brettel, K. ATP Binding Turns Plant Cryptochrome Into an Efficient Natural Photoswitch. *Sci. Rep.* **2014**, *4*, 5175.

(26) Adamo, C.; Jacquemin, D. The calculations of excited-state properties with Time-Dependent Density Functional Theory. *Chem. Soc. Rev.* **2013**, *42*, 845–856.

(27) Aksimentiev, A.; Schulten, K. Imaging alpha-hemolysin with molecular dynamics: Ionic conductance, osmotic permeability and the electrostatic potential map. *Biophys. J.* **2005**, *88*, 3745–3761.

(28) Lüdemann, G.; Solov'yov, I. A.; Kubař, T.; Elstner, M. Solvent driving force ensures fast formation of a persistent and well-separated radical pair in plant cryptochrome. *J. Am. Chem. Soc.* **2015**, *137*, 1147–1156.