Electronic Supplementary Material (ESI) for Chemical Communications.

Electronic Supplementary Information

Discovery and Functional Analysis of a 4th Electron-Transferring Tryptophan Conserved Exclusively in Animal Cryptochromes and (6-4) Photolyases

Pavel Müller,*[a] Junpei Yamamoto,[b] Ryan Martin,[c] Shigenori Iwai,[b] and Klaus Brettel*[a]

[a]Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud; CEA Saclay, iBiTec-S/SB2SM, Bât. 532, 91191 Gif-sur-Yvette, France (E-mail: pavel.muller@cea.fr, klaus.brettel@cea.fr)
[b]Division of Chemistry, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan
[c]Department of Chemistry, UMR 8640 CNRS-ENS-UPMC, Ecole Normale Supérieure, 24 rue Lhomond, 75005 Paris, France

Experimental Details

Sample preparation. The wild type *Xenopus laevis* (6-4) photolyase (*Xl*(6-4)PL) was prepared using a procedure reported previously[1]. For the preparation of the W370F mutant of *Xl*(6-4)PL, a pET28a-based construct encoding His-tagged *Xl*(6-4)PL[2], which was kindly provided by Dr. Hideki Kandori (Nagoya Institute of Technology, Japan), was used as a template, and site-directed mutagenesis was performed with PrimeSTAR® Mutagenesis Basal Kit (TaKaRa, Japan). The set of PCR primers used in the mutagenesis was the following: d(GAC CTC TTC ATA TCA TGG GAA GAA GGA) and d(TGA TAT GAA GAG GTC CCC TCG GGT GAG). After DNA sequencing of the obtained plasmid to confirm the introduction of mutagenesis, the mutated plasmid was transfected into OverExpress C41(DE3) competent cells (Lucigen Corporation, US). The transfected cells were cultured in LB medium with 100 μg/ml of ampicillin at 25°C for 9 hours, and isopropyl β-D-1-thiogalactopyranoside was then added into the medium to a final concentration of 1 mM. The mixture was further shaken at 25°C for 24 hours, and the cells were harvested by centrifugation. Lysozyme (50 mg) and glycerol (2 mL) was added to the harvested pellets, and the mixture was suspended in a lysis buffer (40 mL) containing 50 mM NaH_2PO_4, 200 mM NaCl, and 5 mM imidazole (pH 8). The cells were disrupted by sonication on ice. The cell extract was centrifuged, and the obtained supernatant was loaded on a column containing TALON® Superflow Metal Affinity Resin (TaKaRa, Japan) in a cold chamber (4°C). The resin was washed with the lysis buffer, and the bound enzyme was eluted with an elution buffer containing 50 mM NaH_2PO_4, 200 mM NaCl, and 500 mM imidazole (pH 8). The obtained yellow solution was immediately loaded onto a HiTrap Heparin HP column (GE Healthcare, US), and the elution was performed by an ÄKTA purifier (GE Healthcare) with a buffer containing 50 mM Tris-HCl, 100-600 mM NaCl, and 5% glycerol (pH 8). The fractions containing enzyme were pooled, and the buffer was exchanged into 50 mM

Tris-HCl, 50 mM NaCl, and 5% glycerol (pH 8), by repeated dilution-concentration processes. The concentrated stock solution was frozen in liquid nitrogen and stored at –80°C. Before use, the samples of both wild-type and W370F *Xl*(6-4)PL were rid of possible free FAD using Micro Bio-Spin (Bio-Gel P-6) size-exclusion chromatography columns pre-washed with the same buffer (50 mM Tris-HCl, 50 mM NaCl, and 5% glycerol, pH 8.0 at 20°C).

Spectroscopic experiments. UV-Vis spectra and transient absorption kinetics were measured as described[3], with the following modifications:

For kinetic measurements up to 80 µs (with a 20 MHz bandwidth limit), the monitoring light was provided by the following continuous wave lasers:

376 nm – diode laser Toptica iBeam smart 375-S (up to 120 mW),

408 nm – diode laser Toptica iBeam smart 405-S (up to 120 mW),

448 nm – laser diode Nichia NDHB510 (50 mW),

457 nm – diode pumped solid state (DPSS) laser Cobolt Twist™ (50 mW),

488 nm – diode laser Picarro Cyan-20 from Spectra-Physics (20 mW),

515 nm – DPSS laser Cobolt Fandango™ (150 mW),

562 nm – DPSS laser Oxxius 561-25-COL-002 (25 mW),

594 nm – DPSS laser Cobolt Mambo™ (100 mW),

638 nm – diode laser Toptica iBeam smart 640-S (up to 150 mW).

2×2×10 mm cells were used (excitation pulses entered the sample through the 2×10 mm window, monitoring light through the 2×2 mm window). The monitoring light beams were attenuated by neutral density filters and mechanically chopped to produce light pulses of 140 µs duration and energy in the order of 1 µJ at the entrance of the cell.

For experiments up to 0.4 s (with a 300 Hz bandwidth limit), interference filters with transmission maxima at 383, 410, 450, 460, 490, 520, 560, 590 or 630 nm and spectral bandwidths of 5 to 10 nm were placed between the tungsten-halogen lamp and the sample.

Each kinetic trace results from a single-flash excitation (no signal averaging) of a sample containing >95% FAD_{ox} (checked from UV-Vis spectra). The samples were air-saturated and kept at 10°C during experiments and on ice in between. The pH of the used Tris-HCl buffer was 8.3 at 10°C (8.0 at 20°C).

Signal Analysis. Transient absorption signals were fitted globally with common time constant(s) τ_1 (and τ_2) using Origin 8.6. While signals of the WT protein both at the ns/µs and ms/s timescales (Figs. S3a and S4a, respectively) could be well fitted by monoexponential decay functions

$$\Delta A = A_1\, e^{-\frac{1}{\tau_1}t} + y_0 \qquad (1),$$

bi-exponential decay function had to be used for fitting of the ns/µs signals of the W370F mutant (Fig. S3b):

$$\Delta A = A_1\, e^{-\frac{1}{\tau_1}t} + A_2\, e^{-\frac{1}{\tau_2}t} + y_0 \qquad (2).$$

The time constants obtained in the fits are summarized in Table S1. Initial amplitudes for the Figs. 4b, 4c and S4b were obtained as follows: ΔA for $t \rightarrow 0$ were calculated as A_1 (+ A_2) + y_0. ΔA at t = 3 µs were read directly from the curves resulting from the fit. ΔA at the end of the kinetic phases correspond to y_0.

Quantum yield determination. Using the same concentration, the same excitation energy and the same setup geometry, we observed transient absorption changes of approximately the same initial amplitudes for both WT and W370F mutant proteins (Figs. 3 and S3), indicating similar quantum yields for formation of $FAD^{\bullet -}$ $Trp_4H^{\bullet +}$ (in the WT) and $FAD^{\bullet -}$ $Trp_3H^{\bullet +}$ (in the W370F protein). Since the spectrum of the $TrpH^{\bullet +}$ radical in the WT protein seems to slightly deviate from the published spectra of free $TrpH^{\bullet +}$ cation radicals, we have decided to estimate the quantum yield of the terminal $FAD^{\bullet -}$ $TrpH^{\bullet +}$ pair from the signals of the W370F protein.

We have used an aqueous solution of 25 µM $[Ru(bpy)_3]Cl_2$ (99.95%) as actinometer[4]. Under the same geometry and excitation (by laser pulses at 355 nm) as used for the ns/µs data of the *Xl*(6-4)PL samples, the $[Ru(bpy)_3]^{2+}$ sample showed an initial absorption change of -46.4 mOD at 448 nm over the 1 cm optical path, corresponding to 4.22 µM of excited Ru complexes (calculated with $\Delta\varepsilon_{448}$ (^{3}MLCT – $[Ru(bpy)_3]^{2+}$) = -11 000 $M^{-1}cm^{-1}$ and a quantum yield of the metal-to-ligand charge transfer triplet (^{3}MLCT) of 100%)[4a]. For the 70 µM W370F PL sample, the initial absorption change of -22.3 mOD at 457 nm (where FAD_{ox} bleaching is still pronounced and the signal-to-noise ratio was better than at 448 nm) corresponds to 4.63 µM $FAD^{\bullet -}$ $TrpH^{\bullet +}$ (calculated with ($\Delta\varepsilon_{457}$ = $\varepsilon_{457}(FAD^{\bullet -})$ + $\varepsilon_{457}(TrpH^{\bullet +})$ - $\varepsilon_{457}(FAD_{ox})$ = (4770 + 350 - 9935) $M^{-1}cm^{-1}$ = -4815 $M^{-1}cm^{-1}$). Based on their respective optical densities at 355 nm of 0.145 and 0.0318 (along the 2 mm excitation path), the *Xl*(6-4)PL samples absorbed 4.0 times more of the incident 355 nm photons than the Ru complex. Hence, 4.0 × 4.22 µM = 16.88 µM FAD_{ox} were excited in the *Xl*(6-4)PL samples. The observed 4.63 µM of the terminal $FAD^{\bullet -}$ $TrpH^{\bullet +}$ pair correspond therefore to a quantum yield of 27.4%.

In this estimation, we neglected the possibility of non-productive absorption of a second photon by the same molecule, which would diminish the reaction yield per absorbed photon. As the absorption coefficient (and hence the absorption cross section) of FAD_{ox} at 355 nm is higher than that of the Ru complex (~10 vs. ~6 $mM^{-1}cm^{-1}$), it is likely that we slightly underestimated the quantum yield in the protein. Further considering pulse-to-pulse fluctuations of the excitation laser energy (in the order of 5%) and uncertainties on the exact values of the absorption coefficients used, we estimate the quantum yield of formation of the terminal $FAD^{•-}$ $TrpH^{•+}$ pair to be 30±5%. The ~70% losses (in both WT and W370F proteins) are presumably due to fast (< 1 ns) recombination of the $FAD^{•-}$ $Trp_1H^{•+}$ or $FAD^{•-}$ $Trp_2H^{•+}$ radical pairs that could not be resolved using our setup.

Supplemental Structure and Sequence Analysis

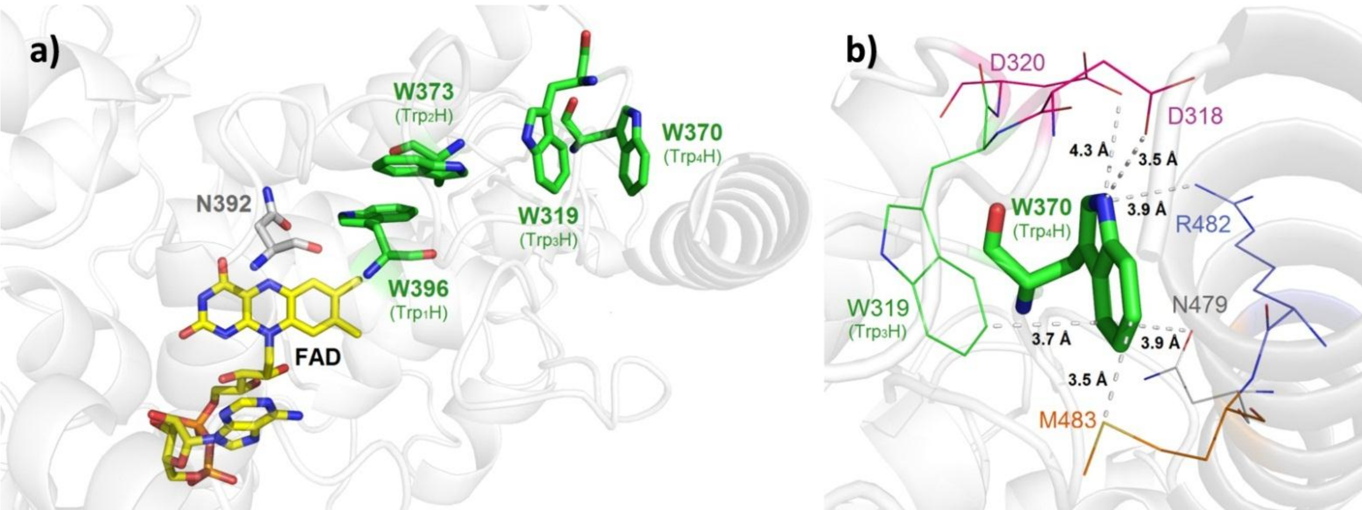

Figure S1. Homology model of the *Xl*(6-4)PL structure. *Xl*(6-4)PL was aligned to the known structure of Drosophila (6-4) PL (structure 3CVU in RCSB PDB)[5] using the SWISS-MODEL platform[6]. (a) FAD is shown in yellow, the tetrad of Trps involved in ET to photoexcited FAD in green. Asparagine N392 (grey) faces the N5 atom of the FAD isoalloxazine (which becomes protonated in the case of $FAD^{•-}$ conversion to $FADH^{•}$).
(b) Detailed picture of the immediate surroundings of the 4th Trp (W370) in *Xl*(6-4)PL. All shown amino acids except D318 are highly conserved among animal Crys and animal (6-4) PLs (see Figure S2).

```
                      D318 D320                                                                Trp4                  FAD
                          Trp3                                                                      Trp2              N5    Trp1
EcCPD    295 HRPFIAWTDR VQWQSN-PAH LQAWQEGKTG YPIVDAAMRQ LNSTGWMHNR LRMITASFLV K-DLLIDWRE GERYFMSQLI DGDLAANNGG WQWAASTGTD 392
TtCPD    263 ERPLDPRFQA LPWQED-EAL FRAWYEGRTG VPLVDAAMRE LHATGFLSNR ARMNAAQFAV K-HLLLPWKR CEEAFRHLLL DGDRAVNLQG WQWAGGLGVD 360
AnCPD    301 DGPYRSLWQQ FPWENR-EAL FTAWTQAQTG YPIVDAAMRQ LTETGWMHNR CRMIVASFLT K-DLIIDWRR GEQFFMQHLV DGDLAANNGG WQWSASSGMD 398
Ds64     354 RIAGNPICRQ ITWDTN-PAL LKAWRDGATG YPWIDAAMTQ LREWGWMHHL ARHSVACFLT RGDLYLSWES GKEVFEELLL DADYFINAAN WMWLSASAFF 452
At64     317 KMKGNRICKQ IPWNED-HAM LAAWRDGKTG YPWIDAIMVQ LLKWGWMHHL ARHCVACFLT RGDLFIHWEQ GRDVFERLLI DSDWAINNGN WMWLSCSSFF 415
Dm64     318 RMLGNVYCMQ IPWQEH-PDH LEAWTHGRTG YPFIDAIMRQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GQRVFEQLLL DQDWALNAGN WMWLSASAFF 416
Dr64     307 KMEGNSACVQ VDWDNN-PEH LAAWREARTG FPFIDTIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GQKVFEELLL DSDWSLNAGN WQWLSASTFF 405
Xl64     307 KMEGNPVCVQ VDWDNN-KEH LEAWSEGRTG YPFIDAIMTQ LRTEGWIHHL ARHAVACFLT RGDLWISWEE GQKVFEELLL DADWSLNAGN WLWLSASAFF 405
AtCRY1   312 ERPLLGHLKF FPWAVD-ENY FKAWRQGRTG YPLVDAGMRE LWATGWLHDR IRVVVSSFFV K-VLQLPWRW GMKYFWDTLL DADLESDALG WQYITGTLPD 409
AtCRY2   307 EQSLLSHLRF FPWDAD-VDK FKAWRQGRTG YPLVDAGMRE LWATGWMHNR IRVIVSSFAV K-FLLLPWKW GMKYFWDTLL DADLECDILG WQYISGSIPD 406
AtCRY3   388 FHLGGPRNVQ GKWSQD-QKL FESWRDAKTG YPLIDANMKE LSTTGFMSNR GRQIVCSFLV R-DMGLDWRM GAEWFETCLL DYDPCSNYGN WTYGAGVGND 485
OtCPF1   339 FHLDGTAGRR ASWKRD-EKI LKAWKTGTTG YPLIDANMRE LAATGFMSNR GRQNVASWLA L-DAGIDWRH GADWFEHHLL DYDTASNWGN WCAAAGMTGG 436
PtCPF1   339 KMIDNPIARQ IPWDDD-PDL LLAWKMSKTG YPYIDAIMTQ LRETGWIHHL ARHSVACFLT RGDLWQSWED GATVFEEYLI DADWSINNFN WQWLSCTAHF 437
DmCRY1   330 RMEGNDICLS IPWAKPNENL LQSWRLGQTG FPLIDGAMRQ LLAEGWLHHT LRNTVATFLT RGGLWQSWEH GLQHFLKYLL DADWSVCAGN WMWVSSSAFE 429
DrCRY1a  308 KMEGNPICVQ IPWDKN-PEA LAKWAEGRTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSVNAGS WMWLSCSSFF 406
DrCRY1b  308 KMEFNPICVQ IPWDRN-PEA LAKWAEGQTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSVNAGS WMWLSCSSFF 406
DrCRY2a  308 KMEGNPICVR IPWDKN-PEA LAKWAEAKTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSVNAGS WMWLSCSSFF 406
DrCRY2b  310 RMEGNPICIR IPWDRN-AEA LAKWAEAKTG FPWIDAIMMQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSVNAGS WLCHSCSSFF 408
DrCRY4   305 KMEGNSICLQ IDWYHD-PER LEKWRTAQTG FPWIDAIMTQ LLQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEEFLL DADYSVNAGN WMWLSASAFF 403
XlCRY1   307 HMVGNPICLQ IEWYKN-EEQ LQKWREGKTG FPWIDAIMAQ LHEEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADYSINAGN WMWLSASAFF 405
XlCRY2   312 QMEGNPICVQ IPWDKN-PKA LAKWTEGKTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWNSWEC GVKVFDELLL DADFSVNAGS WMWLSCSAFF 410
XlCRYD   312 FFLRGLQDKD IPWKRD-PKL FDAWKEGRTG VPFVDANMRE LAMTGFMSNR GRQNVASFLT K-DLGIDWRM GAEWFEYLLV DYDVCSNYGN WLYSAGIGND 409
ErCRY1a  308 KMEGNPICVQ IPWDKN-PEA LAKWAEGRTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSVNAGS WMWLSCSSFF 406
ErCRY1b  308 KMEGNPICVQ IPWDKN-PEA LAKWAEGRTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSVNAGS WMWLSCSSFF 406
GgCRY1   308 KMEGNPICVQ IPWDKN-PEA LAKWAEGRTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSVNAGS WMWLSCSSFF 406
GgCRY2   317 RMEGNPICIQ IPWDKN-PEA LAKWAEGKTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWES GVRVFDELLL DADFSVNAGS WMWLSCSAFF 415
GgCRY4   306 KMAGNPICLQ IRWYED-AER LHKWKTAQTG FPWIDAIMTQ LRQEGWIHHL ARHAAACFLT RGDLWISWEE GMKVFEELLL DADYSINAGN WMWLSASAFF 404
MmCRY1   308 KMEGNPICVQ IPWDKN-PEA LAKWAEGRTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSINAGS WMWLSCSSFF 406
MmCRY2   326 RMEGNPICIQ IPWDRN-PEA LAKWAEGKTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWVSWES GVRVFDELLL DADFSVNAGS WMWLSCSAFF 424
HsCRY1   308 KMEGNPICVQ IPWDKN-PEA LAKWAEGRTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWISWEE GMKVFEELLL DADWSINAGS WMWLSCSSFF 406
HsCRY2   348 RMEGNPICIQ IPWDRN-PEA LAKWAEGKTG FPWIDAIMTQ LRQEGWIHHL ARHAVACFLT RGDLWVSWES GVRVFDELLL DADFSVNAGS WMWLSCSAFF 446
```

```
                                                                                              R482
                                                                                          N479  M483
EcCPD    393 AAPY--FRIF NPTTQGEKFD HEGEFIRQWL PELRDVPGKV VHEPWKWAQK AG------VT LDYPQPIVEH KEARVQTLAA YEAARKGK-- ---------- 472
TtCPD    361 AAPY--FRVF NPVLQGERHD PEGRWLKRWA PEYPSYAPKD PVVD---LEE AR-------- RRYLR----- ---------L ARDLARG--- ---------- 420
AnCPD    399 PKP---LRIF NPASQAKKFD ATATYIKRWL PELRHVHPKD LISG---EIT PI------ER RGYPAPIVNH NLRQKQFKAL YNQLKAAI-- ---------- 474
Ds64     453 -AQY--FRVY SPVVFGKKYD KEGAYIRKFL PVLKDMPAKY IYEPWTAPKE VQQRANCIIG RDYPAPIVDH AVASKECIAR MGAAYKATNT GGSAGKASPA 549
At64     416 -YQF--NRIY SPISFGKKYD PDGKYIRHFL PVLKDMPKQY IYEPWTAPLS VQTKANCIVG KDYPKPMVLH DSASKECKRK MGEAYALNKK MDGKVDEENL 512
Dm64     417 -HQY--FRVY SPVAFGKKTD PQGHYIRKYV PELSKYPAGC IYEPWKASLV DQRAYGCVLG TDYPHRIVKH EVVHKENIKR MGAAYKVNRE VRTGKEEESS 513
Dr64     406 -HQY--FRVY SPIAFGKKTD KHGDYIKKYL PVLKKFPTEY IYEPWKAPRS VQERAGCIVG KDYPRPIVDH EVVHKKNILR MKAAYAKRSP EDKT-----I 497
Xl64     406 -HQF--FRVY SPVAFGKKTD KNGDYIKKYL PILKKFPAEY IYEPWKSPRS LQERAGCIIG KDYPKPIVEH NVVSKQNIQR MKAAYARRSG STEG-----V 497
AtCRY1   410 SREF--DRID NPQFEGYKFD PNGEYVRRWL PELSRLPTDW IHHPWNAPES VLQAAGIELG SNYPLPIVGL DEAKARLHEA LSQMWQLEAA SRAA-----I 502
AtCRY2   407 GHEL--DRLD NPALQGAKYD PEGEYIRQWL PELARLPTEW IHHPWDAPLT VLKASGVELG TNYAKPIVDI DTARELLAKA ISRTREAQIM IGAAPDE--I 502
AtCRY3   486 PRE---DRYF SIPKQAQNYD PEGEYVAFWL QQLRRLPKEK RHWPGRLMYM DT-------- -VVPLKHGNG PM---AGGSK SGGGFRGS-- ---------- 558
OtCPF1   437 --R---INRF NIAKQTKDYD PAGEYIKTWV KELAEVPAAY IADPNQAPRE LR----DRIG LNYPNKLALP RRDFTEMGSP -PGPRRGGGG GGR------- 519
PtCPF1   438 -YQY--FRCY SPIAFGKKTD PNGDYIRKWL PQFKDMPAKY IYEPWEAPIE LQKKVGVIVG ENYPHPIVDH KLVSKNNMSR MKEAYDAQKN REPMPANESH 534
DmCRY1   430 -RLLDSSLVT CPVALAKRLD PDGTYIKQYV PELMNVPKEF VHEPWRMSAE QQEQYECLIG VHYPERIIDL SMAVKRNMLA MKSLRNSLIT PPPHCRPSNE 528
DrCRY1a  407 -QQF--FHCY CPVSFGRRTD PNGDYIRRYL PVLRGFPAKY IYDPWNAPES VQKAAKCIIG VHYPMPMVHH AEASRLNIER MKQIYQQLSC YRGLGLLAMV 503
DrCRY1b  407 -QQF--FHCY CPVGFGRRTD PNGDYIRRYL PILRGFPAKF IYDPWNAPES VQKVAKCIIG VHYPKPMVNH AEASRINIER MKQIYQQLSC YRGLGLLATI 503
DrCRY2a  407 -QQF--FHCY CPVGFGRRTD PNGDFIRRYL PILRGFPAKY IYDPWNAPDS VQAAAKCIIG VHYPKPMVNH AEASRLNIER MKQIYQQLSR YRGLGLLASV 503
DrCRY2b  409 -QQF--FHCY CPVGFGRRID PNGDFIRRYL PVLRDFPAKY IYDPWNAPHD VQLAAKCAIG VDYPKPMVNH AEASRLNIER MRQIYQQLSR YRGLSLLATV 505
DrCRY4   404 -HKY--TRIF CPVRFGRRTD PQGEYLRKYL PVLKNFPSQY IYEPWKAPED VQLSAGCIIG KDYPRPIVSH IEASQRNLEL MRQVRTEQQT TAELTRD--V 498
XlCRY1   406 -HHY--TRIF CPVRFGRRTD PEGNYIRKYL PVLKNFPAKY IYAPWTAPEE IQKQSGCLIG KDYPLPMVDH TTASEHNLQL MRLVREAQQK TAQLTTD--I 500
XlCRY2   411 -QQF--FHCY CPVGFGRRTD PSGDYVKRYL PVLKAFPSRY IYEPWSAPES VQKEAKCIIG IDYPKPIVNH AEASRMNIER MKQTYQQLSH YRGLCILASV 507
XlCRYD   410 PRE---NRKF NMIKQGLDYD SGGDYIRLWV PELQQIKGGD AHTPWALSNA SLAHANLSLG ETYPYPIVMA PEWSRHINQK PAGSWEKSAR RGK------- 499
ErCRY1a  407 -QQF--FHCY CPVGFGRRTD PNGDYIRRYL PVLRGFPAKY IYDPWNAPES IQKAAKCIIG VNYPKPMVNH AEASRLNIER MKQIYQQLSR YRGLGLLATV 503
ErCRY1b  407 -QQF--FHCY CPVGFGRRTD PNGDYIRRYL PILRGFPAKY IYDPWNAPES IQKAAKCIIG VNYPKPMVNH AEASRLNIER MKQIYQQLSR YRGLGLLATV 503
GgCRY1   407 -QQF--FHCY CPVGFGRRTD PNGDYIRRYL PVLRGFPAKY IYDPWNAPES VQKAAKCVIG VNYPKPMVNH AEASRLNIER MKQIYQQLSR YRGLGLLATV 503
GgCRY2   416 -QQF--FHCY CPVGFGRRTD PSGDYVKRYL PKLKGFPSRY IYEPWNAPES VQKAAKCIIG VDYPKPMVNH AETSRLNIER MKQIYQQLSR YRGLCLLASV 512
GgCRY4   405 -HHY--TRIF CPVRFGRRTD PEGQYIRKYL PILKNFPSKY IYEPWTASEE EQKQAGCIIG RDYPFPMVDH KEASDHNLQL MKQAREEQHR IAQLTRD--D 499
MmCRY1   407 -QQF--FHCY CPVGFGRRTD PNGDYIRRYL PVLRGFPAKY IYDPWNAPEG IQKVAKCLIG VNYPKPMVNH AEASRLNIER MKQIYQQLSR YRGLGLLASV 503
MmCRY2   425 -QQF--FHCY CPVGFGRRTD PSGDYIRRYL PKLKGFPSRY IYEPWNAPES VQKAAKCIIG VDYPRPIVNH AETSRLNIER MKQIYQQLSR YRGLCLLASV 521
HsCRY1   407 -QQF--FHCY CPVGFGRRTD PNGDYIRRYL PVLRGFPAKY IYDPWNAPEG IQKVAKCLIG VNYPKPMVNH AEASRLNIER MKQIYQQLSR YRGLGLLASV 503
HsCRY2   447 -QQF--FHCY CPVGFGRRTD PSGDYIRRYL PKLKAFPSRY IYEPWNAPES IQKAAKCIIG VDYPRPIVNH AETSRLNIER MKQIYQQLSR YRGLCLLASV 543
```

Figure S2. Sequence alignments of Cry/PL proteins. Ends of the Photolyase Homology Region (PHR) of several representatives of various Cry/PL protein classes are shown, highlighting the tryptophan triad (red on grey background) conserved in all proteins and the fourth tryptophan (white on red background) occurring only in animal Crys, animal (6-4) photolyases and also in the dual protein *Pt*CPF1 from *Phaeodactylum tricornutum*. Proteins containing the 4th tryptophan have also relatively well conserved amino acids in its nearest vicinity (see Figure S1): methionine (pink background, at the position of M483 in *Xl*(6-4)PL), arginine/lysine (dark green background, at the position of R482 in *Xl*(6-4)PL), asparagine (blue background, at the position of N479 in *Xl*(6-4)PL), and aspartate/glutamate (orange background, at the position of D320 in *Xl*(6-4)PL). The second aspartate close to the 4th tryptophan in *Xl*(6-4)PL (D318) is conserved only in a few (6-4) PLs and lower vertebrate cryptochromes, other species contain mostly proline at this position. The asparagine facing the N5 atom of the FAD isoalloxazine in most Cry/PL proteins is shown with light green background; insect type I Crys (*Dm*CRY1) contain cysteine at this position and plant Crys have an aspartic acid there.

Abbreviations:

Ec = Escherichia coli	*Tt = Thermus thermophilus*	*An = Anacystis nidulans*	*Ds = Dunaliella salina*	*At = Arabidopsis thaliana*
Dm = Drosophila melanogaster	*Dr = Danio rerio*	*Xl = Xenopus laevis*	*Ot = Ostreococcus tauri*	*Pt = Phaeodactylum tricornutum*
Er = Erithacus rubecula	*Gg = Gallus gallus*	*Mm = Mus musculus*	*Hs = Homo sapiens*	

CPD = Cyclobutane Pyrimidine Dimer photolyase

64 = (6-4) photolyase

CRY = Cryptochrome

CRYD = “*DASH*” cryptochrome

CPF = Cryptochrome Photolyase Family

Supplemental Transient Absorption Data

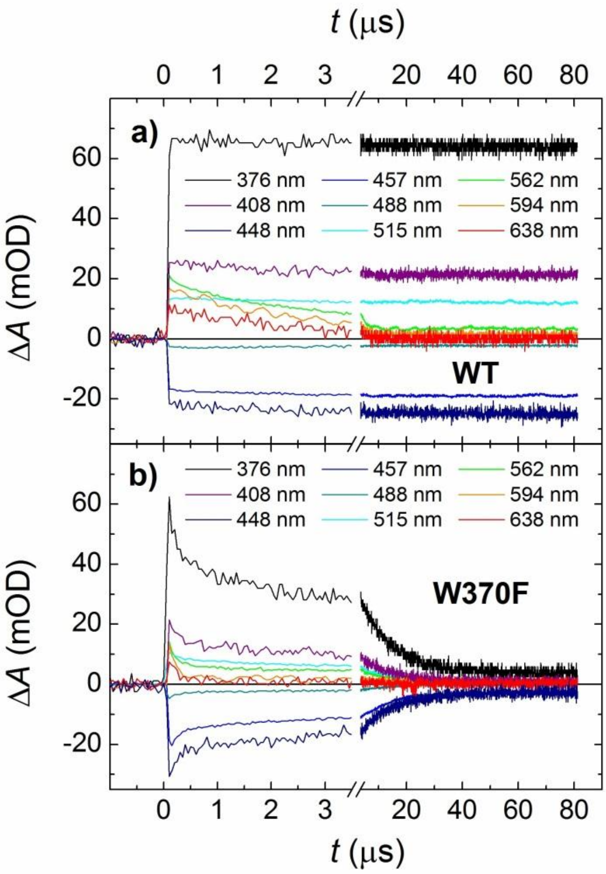

Figure S3. Nano- and microsecond transient absorption kinetics. Flash-induced absorption changes were recorded for (a) WT *Xl*(6-4)PL and (b) its W370F mutant at the indicated wavelengths. FAD_{ox} was excited by a laser pulse at $t = 0$ (355 nm, 100 ps pulse duration, E ~4.0 mJ.cm^{-2}. The samples were kept at 10°C. They contained 70 µM protein in 50 mM Tris-HCl buffer of pH 8.3 (at 10°C), 50 mM NaCl and 5% glycerol.

Protein	**Timescale** (Fitted data)	τ_1	τ_2
WT	**ns/µs** (Figure S3a)	**(2.556 ± 0.009) × 10^{-6} s**	**-**
WT	**ms/s** (Figure S4a)	**(3.423 ± 0.005) × 10^{-2} s**	**-**
W370F	**ns/µs** (Figure S3b)	**(1.886 ± 0.015) × 10^{-7} s**	**(1.052 ± 0.003) × 10^{-5} s**

Table S1. Fit results. Time constants τ_1 and τ_2 were obtained by global fitting of data in Figs. S3 and S4 by mono- (WT protein) and bi-exponential (W370F mutant) decay functions: equations (1) and (2) on page S3.

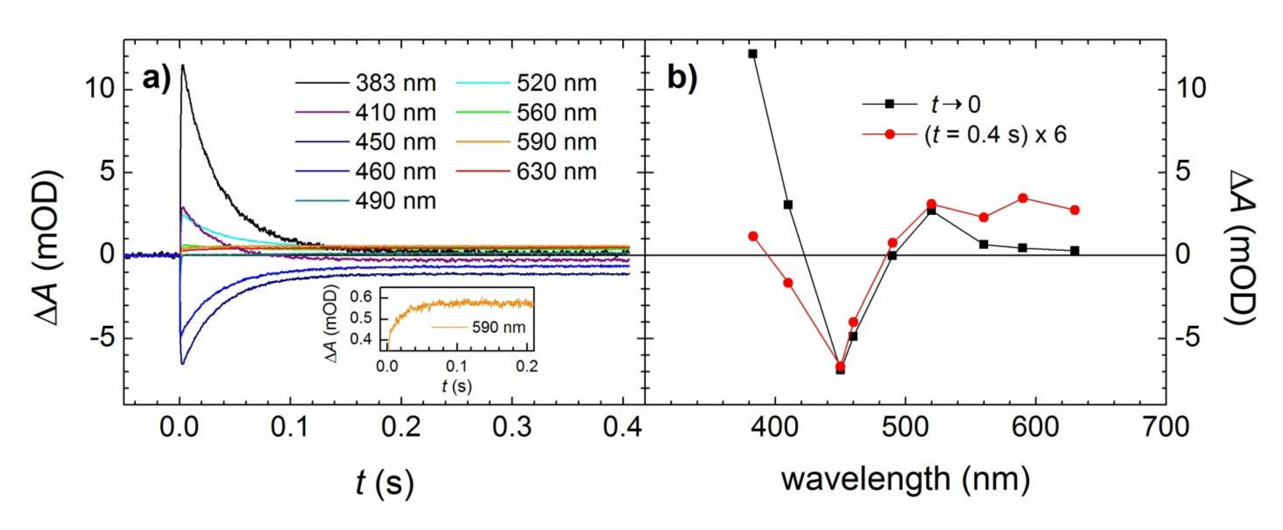

Figure S4. Millisecond transient absorption kinetics. (a) Flash-induced absorption changes in WT *Xl*(6-4)PL (20 µM protein in 50 mM Tris-HCl buffer at pH 8.3 [at 10°C], 50 mM NaCl and 5% glycerol) were recorded at the indicated wavelengths. FAD_{ox} was excited by a laser pulse at $t = 0$ (355 nm, 100 ps duration, E ~4.0 mJ.cm^{-2}). Global fit of the decay curves yields a time constant τ ~35 ms (see Table S1). Inset: Expansion of the 590 nm trace showing an absorption increase with τ ~35 ms attributed to formation of the neutral $FADH^{\bullet}$ radical. (b) Amplitudes of the kinetic traces shown in a), extrapolated to $t \rightarrow 0$ according to the global fit (black squares) and amplitudes at $t = 0.4$ s (red circles, multiplied by a factor of 6 for a better qualitative comparison with the points at $t \rightarrow 0$).

Supplemental Discussion

Particularities of the absorption spectra of Trp_4 radicals and the deprotonation kinetics of $Trp_4H^{\bullet+}$.

According to the published spectra of (free) $TrpH^{\bullet+}$ and $Trp^{\bullet}$, deprotonation of $TrpH^{\bullet+}$ should lead to a slight increase of absorbance around 450 nm but we have observed the opposite trend in the WT protein (slight absorption decrease – Figures 3a and S3a). Furthermore, in contrast to the W370F mutant protein, in the WT protein, there is a slight positive deviation around 450 nm between the observed and the expected difference spectra for formation of the $FAD^{\bullet-}$ $TrpH^{\bullet+}$ and the $FAD^{\bullet-}$ $Trp^{\bullet}$ pair (Figure 4). These observations indicate that the environment of the 4th tryptophan (W370) has a substantial impact on the spectra of the tryptophanyl radicals. There are several charged, polar or polarizable functional groups in the immediate vicinity of W370 (see Figure S1b), which might affect the electronic states of its radicals: COO^- groups of D318 and D320, $C(NH_2)=NH_2^+$ group of R482, the electronegative oxygen of the N479 amide, or the polarizable sulphur atom of M483. *E.g.*, one could imagine that charge transfer (CT) transitions due to the interactions with M483 enhance the absorbances of $TrpH^{\bullet+}$ and $Trp^{\bullet}$ around 450 nm (*cf.*, methionine to Cu^{2+} CT transitions in plastocyanin and azurin)[7].

The overall negatively charged, polar and polarizable environment of W370 is further expected to stabilize the positively charged $Trp_4H^{\bullet+}$ relative to neutral Trp_4H and $Trp_4^{\bullet}$. This would correspond to an unusually low reduction potential of $Trp_4H^{\bullet+}/Trp_4H$ (that would favour ET from Trp_4H to $Trp_3H^{\bullet+}$) and to an unusually high pK_a value of $Trp_4H^{\bullet+}$ that might explain the rather slow deprotonation of $Trp_4H^{\bullet+}$ in the WT protein (2.5 µs vs. 200 – 400 ns for

$Trp_3H^{•+}$ in the W370F mutant protein (see main text), in *E.coli* PL[8] and in *At*Cry1[3a]). A similarly slow $TrpH^{•+}$ deprotonation (2.56 µs) has been reported very recently for *Drosophila melanogaster* cryptochrome.[9] Interestingly, *Dm*Cry also possesses the 4th Trp (W394 in *Dm*Cry; see Figures 2 and S2). We consider it very likely that W394 (and not the 3rd Trp W342 as assumed in Ref. [9]) functions as the terminal member of the Trp chain in *Dm*Cry. In line with our model, the $FAD^{•-}$ $Trp^{•}$ radical pair observed after deprotonation of $TrpH^{•+}$ recombined rather slowly in *Dm*CRY (6.8 ms at pH 7.0).[9] In bird (garden warbler) CRY1, the $FAD^{•-}$ $Trp^{•}$ pair was reported to recombine in 14 ms (at pH 7.4).[10] The deviations from 40 ms observed by us in *Xl*(6-4)PL at pH 8.3 are likely due to the different pH values, as low pH accelerates such recombinations.[8, 11] These observations support our initial assumption that the specific features of the flavin – tryptophan radical pairs observed in our study on *Xl*(6-4)PL can be generalized to all proteins with the Trp tetrad including animal cryptochromes.

Competition between recombination of $FAD^{•-}$ $Trp^{•}$* and *$FAD^{•-}$ protonation. The signals on the 0.4 s timescale (Figure S4a) are dominated by decay consistent with recombination of an $FAD^{•-}$ $Trp^{•}$ radical pair but they do not decay completely. The very small absorption changes remaining after the 35 ms kinetic phase show a difference spectrum with a relatively pronounced absorption increase between 520 and 630 nm and an only very weak absorption increase at 383 nm (compared to the bleaching at 450 nm), as expected for an $FADH^{•}$ $Trp^{•}$ pair. The $FADH^{•}$ absorption at 590 nm (inset in Figure S4a) built up with the same kinetics as the decay of $FAD^{•-}$ $Trp^{•}$. Within a simplified parallel reaction scheme, the initial to final amplitude ratio of 6 to 1 at 450 nm (note that $FAD^{•-}$ and $FADH^{•}$ absorb about equally at 450 nm (see Figure 4a) yields intrinsic time constants of ~40 ms for recombination of $FAD^{•-}$ $Trp^{•}$ and ~200 ms for protonation of $FAD^{•-}$. Noteworthy, protonation of $FAD^{•-}$ is about 5 orders of magnitude faster in plant Crys that contain an aspartic acid (D396 in *At*Cry1) at the position of asparagine N392 (facing the N5 atom of the FAD cofactor) in *Xl*(6-4)PL (see Figure S1). This very strong kinetic difference supports previous suggestions that the aspartic acid facing the N5 atom of the FAD is the direct proton donor to $FAD^{•-}$ in plant Crys.[3a, 12]

References

1 J. Yamamoto, R. Martin, S. Iwai, P. Plaza and K. Brettel, *Angew. Chem. Int. Ed.*, 2013, **52**, 7432.

2 D. Yamada, Y. Zhang, T. Iwata, K. Hitomi, E. D. Getzoff and H. Kandori, *Biochemistry*, 2012, **51**, 5774.

3 (a) P. Müller, J.-P. Bouly, K. Hitomi, V. Balland, E. D. Getzoff, T. Ritz and K. Brettel, *Sci. Rep.*, 2014, **4**, 5175; (b) M. Byrdin, V. Thiagarajan, S. Villette, A. Espagne and K. Brettel, *Rev. Sci. Instrum.*, 2009, **80**, 043102.

4 (a) P. Müller and K. Brettel, *Photochem. Photobiol. Sci.*, 2012, **11**, 632; (b) V. Thiagarajan, M. Byrdin, A. P. M. Eker, P. Müller and K. Brettel, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 9402.

5 M. J. Maul, T. R. M. Barends, A. F. Glas, M. J. Cryle, T. Domratcheva, S. Schneider, I. Schlichting and T. Carell, *Angew. Chem. Int. Ed.*, 2008, **47**, 10076.

6 L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey and T. Schwede, *Nat. Protocols*, 2008, **4**, 1.

7 D. R. McMillin and M. C. Morris, *Proc. Natl. Acad. Sci. USA*, 1981, **78**, 6567.

8 C. Aubert, M. H. Vos, P. Mathis, A. P. M. Eker and K. Brettel, *Nature*, 2000, **405**, 586.

9 B. Paulus, C. Bajzath, F. Melin, L. Heidinger, V. Kromm, C. Herkersdorf, U. Benz, L. Mann, P. Stehle, P. Hellwig, S. Weber and E. Schleicher, *FEBS Journal*, 2015, **282**, 3175.

10 M. Liedvogel, K. Maeda, K. Henbest, E. Schleicher, T. Simon, C. R. Timmel, P. J. Hore and H. Mouritsen, *Plos One*, 2007, **2**, 7.

11 M. Byrdin, V. Sartor, A. P. M. Eker, M. H. Vos, C. Aubert, K. Brettel and P. Mathis, *Biochim. Biophys. Acta - Bioenergetics*, 2004, **1655**, 64.

12 T. Langenbacher, D. Immeln, B. Dick and T. Kottke, *J. Am. Chem. Soc.*, 2009, **131**, 14274.