# **Electronic Supplementary Information**

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

Pavel Müller,\*<sup>a</sup> Junpei Yamamoto,<sup>b</sup> Ryan Martin,<sup>c</sup> Shigenori Iwai,<sup>b</sup> and Klaus Brettel\*<sup>a</sup>

<sup>a</sup>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: <u>pavel.muller@cea.fr</u>, <u>klaus.brettel@cea.fr</u>) <sup>b</sup>Division of Chemistry, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

<sup>c</sup>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 (XI(6-4)PL) was prepared using a procedure reported previously<sup>1</sup>. For the preparation of the W370F mutant of X/(6-4)PL, a pET28a-based construct encoding His-tagged X/(6-4)PL<sup>2</sup>, which was kindly provided by Dr. Hideki Kandori (Nagova 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<sub>2</sub>PO<sub>4</sub>, 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<sub>2</sub>PO<sub>4</sub>, 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 *XI*(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<sup>3</sup>, 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<sup>™</sup> (50 mW),
- 488 nm diode laser Picarro Cyan-20 from Spectra-Physics (20 mW),
- 515 nm DPSS laser Cobolt Fandango<sup>™</sup> (150 mW),
- 562 nm DPSS laser Oxxius 561-25-COL-002 (25 mW),
- 594 nm DPSS laser Cobolt Mambo<sup>™</sup> (100 mW),

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

 $2\times2\times10$  mm cells were used (excitation pulses entered the sample through the  $2\times10$  mm window, monitoring light through the  $2\times2$  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).

S2

*Signal Analysis.* Transient absorption signals were fitted globally with common time constant(s)  $\tau_1$  (and  $\tau_2$ ) 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

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

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

$$\Delta A = A_1 e^{-\frac{1}{\tau_1}t} + A_2 e^{-\frac{1}{\tau_2}t} + y_0$$
(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:  $\Delta A$  for  $t \rightarrow 0$  were calculated as  $A_1 (+ A_2) + y_0$ .  $\Delta A$  at  $t = 3 \ \mu$ s were read directly from the curves resulting from the fit.  $\Delta A$  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<sup>--</sup> Trp<sub>4</sub>H<sup>++</sup> (in the WT) and FAD<sup>+-</sup> Trp<sub>3</sub>H<sup>++</sup> (in the W370F protein). Since the spectrum of the TrpH<sup>++</sup> radical in the WT protein seems to slightly deviate from the published spectra of free TrpH<sup>++</sup> cation radicals, we have decided to estimate the quantum yield of the terminal FAD<sup>+-</sup> TrpH<sup>++</sup> pair from the signals of the W370F protein.

We have used an aqueous solution of 25  $\mu$ M [Ru(bpy)<sub>3</sub>]Cl<sub>2</sub> (99.95%) as actinometer<sup>4</sup>. Under the same geometry and excitation (by laser pulses at 355 nm) as used for the ns/µs data of the *XI*(6-4)PL samples, the [Ru(bpy)<sub>3</sub>]<sup>2+</sup> sample showed an initial absorption change of -46.4 mOD at 448 nm over the 1 cm optical path, corresponding to 4.22  $\mu$ M of excited Ru complexes (calculated with  $\Delta \varepsilon_{448}$  (<sup>3</sup>MLCT – [Ru(bpy)<sub>3</sub>]<sup>2+</sup>) = -11 000 M<sup>-1</sup>cm<sup>-1</sup> and a quantum yield of the metal-to-ligand charge transfer triplet (<sup>3</sup>MLCT) of 100%)<sup>4a</sup>. For the 70  $\mu$ M W370F PL sample, the initial absorption change of -22.3 mOD at 457 nm (where FAD<sub>ox</sub> bleaching is still pronounced and the signal-to-noise ratio was better than at 448 nm) corresponds to 4.63  $\mu$ M FAD<sup>--</sup> TrpH<sup>++</sup> (calculated with ( $\Delta \varepsilon_{457} = \varepsilon_{457}$ (FAD<sup>--</sup>) +  $\varepsilon_{457}$ (TrpH<sup>++</sup>) -  $\varepsilon_{457}$ (FAD<sub>ox</sub>) = (4770 + 350 - 9935) M<sup>-1</sup>cm<sup>-1</sup> = -4815 M<sup>-1</sup>cm<sup>-1</sup>). Based on their respective optical densities at 355 nm of 0.145 and 0.0318 (along the 2 mm excitation path), the *XI*(6-4)PL samples absorbed 4.0 times more of the incident 355 nm photons than the Ru complex. Hence, 4.0 × 4.22  $\mu$ M = 16.88  $\mu$ M FAD<sub>ox</sub> were excited in the *XI*(6-4)PL samples. The observed 4.63  $\mu$ M of the terminal FAD<sup>+-</sup> TrpH<sup>++</sup> 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<sub>ox</sub> at 355 nm is higher than that of the Ru complex (~10 vs. ~6 mM<sup>-1</sup>cm<sup>-1</sup>), 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<sup>--</sup> TrpH<sup>++</sup> pair to be  $30\pm5\%$ . The ~70% losses (in both WT and W370F proteins) are presumably due to fast (< 1 ns) recombination of the FAD<sup>+-</sup> Trp<sub>1</sub>H<sup>++</sup> or FAD<sup>+-</sup> Trp<sub>2</sub>H<sup>++</sup> radical pairs that could not be resolved using our setup.

# **Supplemental Structure and Sequence Analysis**



**Figure S1.** Homology model of the *XI*(6-4)PL structure. *XI*(6-4)PL was aligned to the known structure of Drosophila (6-4) PL (structure 3CVU in RCSB PDB)<sup>5</sup> using the SWISS-MODEL platform<sup>6</sup>. (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<sup>--</sup> conversion to FADH<sup>+</sup>). (b) Detailed picture of the immediate surroundings of the 4<sup>th</sup> Trp (W370) in *XI*(6-4)PL. All shown amino acids except D318 are highly conserved among animal Crys and animal (6-4) PLs (see Figure S2).

		D	318 D320					Trp4		FAD	Free	
FoCPD	205	UDDFT AWTDD		LOVMORCKAC	VDTUDAAMBO	TNETCUMUND	TDMTTACETV		CERVENSOL T			202
T+CPD	263	ERPLOPREOA	LPWOED_EAL	FRAWVEGRTG	VPLVDAAMRE	LHATGELSNR	ARMNAAOFAV	K-HLLLPWKR	CEEAFRHLLL	DGDRAVNLOG	WOWAGGLGVD	360
AnCPD	301	DGPYRSLWOO	FPWENR-EAL	FTAWTOAOTG	VPIVDAAMRO	LTETGWMHNR	CRMIVASFLT	K-DLTTDWRR	GEOFFMOHLV	DGDLAANNGG	WOWSASSGMD	398
Ds64	354	RIAGNPICRO	ITWDTN-PAL	LKAWRDGATG	YPWIDAAMTO	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	RGDIMISWEE	GORVFEQLLL	dqdwal <mark>n</mark> agn	WMWLSASAFF	416
Dr64	307	KMEGNSACVQ	VDWDNN-PEH	LAAWREARTG	FPFIDTIMTQ	LRQEGWIHHL	ARHAVACFLT	RGDLWISWEE	GQKVFEELLL	DSDWSL <mark>N</mark> AGN	WQWLSASTFF	405
X164	307	KMEGNPVCVQ	V <mark>DWD</mark> NN-KEH	LEAWSEGRTG	YPFIDAIMTQ	LRTEGWIHHL	ARHAVACFLT	RGDLWISWEE	GQKVFEELLL	DADWSL <mark>N</mark> AGN	WLWLSASAFF	405
AtCRY1	312	ERPLLGHLKF	FPWAVD-ENY	FKAWRQGRTG	YPLVDAGMRE	LWATGWLHDR	IRVVVSSFFV	K-VLQLPWRW	GMKYFWDTLL	DADLESDALG	WQYITGTLPD	409
AtCRY2	307	EQSLLSHLRF	FP <mark>WD</mark> AD-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	RMIDNPIARQ	IPWDDD-PDL	LLAWKMSKTG	YPYIDAIMTQ	LRETGWIHHL	ARHSVACFLT	RGDLWQSWED	GATVFEEYLI	DADWSINNFN	WQWLSCTAHF	437
DECRYIS	300	KMEGNDICLS	IPWARPNENL IDWDKN_DEA	LOSWRLGQIG	FPLIDGAMRQ	LLAEGWLHHT	ARNIVAILLI	RGGLWUSWER	GLQHFLKILL	DADWSVCAGN	WEWV555AFE	429
DrCRV1h	308	KMEENPICVQ	TPWDRN_PEA	LAKWAEGRIG	FPWIDAIMIQ	LROEGWIHHL	ARHAVACELT	RODINISWEE	GMKVFEELLL	DADWSVNAGS	MMWLSCSSFF	400
DrCRY2a	308	KMEGNPICVR	TPWDKN-PEA	LAKWAEAKTG	FPWIDAIMTO	LROEGWIHHL	ARHAVACELT	RGDIWISWEE	GMKVFEELLL	DADWSVNAGS	WMWLSCSSFF	406
DrCRY2b	310	RMEGNPICIR	IPWDRN-AEA	LAKWAEAKTG	FPWIDAIMMO	LROEGWIHHL	ARHAVACELT	RGDLWISWEE	GMKVFEELLL	DADWSVNAGS	WLCHSCSSFF	408
DrCRY4	305	KMEGNSICLQ	IDWYHD-PER	LEKWRTAQTG	FPWIDAIMTQ	LLQEGWIHHL	ARHAVACFLT	RGDLWISWEE	GMKVFEEFLL	DADYSVNAGN	WMWLSASAFF	403
X1CRY1	307	HMVGNPICLQ	IEWYKN-EEQ	LQKWREGKTG	FPWIDAIMAQ	LHEEGWIHHL	ARHAVACFLT	RGDLWISWEE	GMKVFEELLL	DADYSI <mark>N</mark> AGN	WMWLSASAFF	405
X1CRY2	312	QMEGNPICVQ	IPWDKN-PKA	LAKWTEGKTG	FPWIDAIMTQ	LRQEGWIHHL	ARHAVACFLT	RGDLWNSWEC	GVKVFDELLL	DADFSV <mark>N</mark> AGS	WMWLSCSAFF	410
X1CRYD	312	FFLRGLQDKD	IPWKRD-PKL	FDAWKEGRTG	VPFVDANMRE	LAMTGFMSNR	GRQNVASFLT	K-DLGIDWRM	GAEWFEYLLV	DYDVCS <mark>N</mark> YGN	WLYSAGIGND	409
ErCRY1a	308	KMEGNPICVQ	IPW <mark>D</mark> KN-PEA	LAKWAEGRTG	FPWIDAIMTQ	LRQEGWIHHL	ARHAVACFLT	RGDL <mark>W</mark> ISWEE	GMKVFEELLL	DADWSV <mark>N</mark> AGS	WMWLSCSSFF	406
ErCRY1b	308	KMEGNPICVQ	IP <mark>WD</mark> KN-PEA	LAKWAEGRTG	FPWIDAIMTQ	LRQEGWIHHL	ARHAVACFLT	RGDLWISWEE	GMKVFEELLL	DADWSV <mark>N</mark> AGS	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	RMEGNPICVQ	IPWDKN-PEA	LAKWAEGRTG	FPWIDAIMTQ	LRQEGWIHHL	ARHAVACFLT	RGDIWISWEE	GMKVFEELLL	DADWSINAGS	WMWLSCSSFF	406
HeCRV1	308	KMEGNPICIQ	TPWDKN-PEA	LAKWAEGKIG	FPWIDAIMIQ	LROEGWIHHL	ARHAVACELT	RGDIWUSWES	GVKVFDELLL	DADESVNAGS	WMWLSCSAFF	424
HSCRY2	348	RMEGNPICIO	TPWDRN-PEA	LAKWAEGKTG	FPWIDAIMTO	LROEGWIHHL	ARHAVACELT	RGDIWVSWES	GVRVFDELLL	DADESVNAGS	WMWLSCSAFF	446
FaceD	202	λλονροτς	NDWWOCEVED	HECEFTRONI	DELBURCKU	VHEDWEWAOE	۵ <i>С</i>	IDVBODIUEU	R48 N479 N	2 1483		472
EcCPD	393	AAPYFRIF	NPTTQGEKFD	HEGEFIRQWL	PELRDVPGKV	VHEPWKWAQK	AGVT	LDYPQPIVEH	R48 N479 M KEARVQTLAA	2 1483 YEAARKGK		472
EcCPD TtCPD AnCPD	393 361 399	AAPYFRIF AAPYFRVF PKPLRIF	NPTTQGEKFD NPVLQGERHD NPASOAKKFD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWI.	PELRDVPGKV PEYPSYAPKD PELRHVHPKD	VHEPWKWAQK PVVDLEE LISGEIT	AGVT ARER	LDYPQPIVEH RRYLR RGYPAPIVNH	R48 N479 M KEARVQTLAA	2 1483 YEAARKGK ARDLARG YNOLKAAI		472 420 474
EcCPD TtCPD AnCPD Ds64	393 361 399 453	AAPYFRIF AAPYFRVF PKPLRIF -AOYFRVY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVVFGKKYD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL	PELRDVPGKV PEYPSYAPKD PELRHVHPKD PVLKDMPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE	AGVT AR PIER VOORANCIIG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH	R48 N479 M KEARVQTLAA L NLRQKQFKAL AVASKECIAR	2 483 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT	GGSAGKASPA	472 420 474 549
EcCPD TtCPD AnCPD Ds64 At64	393 361 399 453 416	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -YQFNRIY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVVFGKKYD SPISFGKKYD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRHFL	PELRDVPGKV PEYPSYAPKD PELRHVHPKD PVLKDMPAKY PVLKDMPKQY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWTAPLS	AGVT AR PIER VQQRANCIIG VQTKANCIVG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH KDYPKPMVLH	R48 N479 M KEARVQTLAA L NLRQKQFKAL AVASKECIAR DSASKECKRK	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGEAYALNKK	GGSAGKASPA MDGKVDEENL	472 420 474 549 512
EcCPD TtCPD AnCPD Ds64 At64 Dm64	393 361 399 453 416 417	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -YQFNRIY -HQYFRVY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVVFGKKYD SPISFGKKYD SPVAFGKKTD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRHFL PQGHYIRKYV	PELRDVPGKV PEYPSYAPKD PELRHVHPKD PVLKDMPAKY PVLKDMPKQY PELSKYPAGC	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWTAPLS IYEPWKASLV	AGVT ARP PIER VQQRANCIIG VQTKANCIVG DQRAYGCVLG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH KDYPKPMVLH TDYPHRIVKH	R48 N479 M KEARVQTLAA L NLRQKQFKAL AVASKECIAR DSASKECKRK EVVHKENIKR	2 4483 YEAARKGK YNQLKAAI MGAAYKATNT MGEAYALNKK MGAAYKVNRE	GGSAGKASPA MDGKVDEENL VRTGKEEESS	472 420 474 549 512 513
EcCPD TtCPD AnCPD Ds64 At64 Dm64 Dr64	393 361 399 453 416 417 406	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -YQFNRIY -HQYFRVY -HQYFRVY	NPTTQGEKFD NPVLQGERHD SPVVFGKKYD SPISFGKKYD SPIAFGKKTD SPIAFGKKTD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKIL PDGKYIRHFL PQGHYIRKYV KHGDYIKKYL	PELRDVPGKV PEYPSYAPKD PELRHVHPKD PVLKDMPAKY PVLKDMPKQY PELSKYPAGC PVLKKFPTEY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWTAPLS IYEPWKASLV IYEPWKAPRS	AGVT ARPI VQQRANCIIG VQTKANCIVG DQRAYGCVLG VQERAGCIVG	LDYPQPIVEH RRYLR RGYPAPIVNH KDYPAPIVNH KDYPKPMVLH TDYPHRIVKH KDYPRPIVDH	R48 N479 N KEARVQTLAA L NLRQKQFKAL AVASKECIAR DSASKECKRK EVVHKENIKR EVVHKKNILR	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGEAYALNKK MGAAYKVNRE MKAAYAKRSP	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI	472 420 474 549 512 513 497
EcCPD TtCPD AnCPD Ds64 At64 Dm64 Dr64 X164	393 361 399 453 416 417 406 406	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -YQFNRIY -HQYFRVY -HQYFRVY -HQFFRVY	NPTTQGEKFD NPVLQGERHD SPVVFGKKTD SPISFGKKTD SPVAFGKKTD SPVAFGKKTD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRHFL PQGHYIRKYV KHGDYIKKYL KNGDYIKKYL	PELRDVPGKV PEYPSYAPKD PELRHVHPKD PVLKDMPAKY PVLKDMPKQY PELSKYPAGC PVLKKFPTEY PILKKFPAEY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWTAPLS IYEPWKASLV IYEPWKAPRS IYEPWKSPRS	AGVT ARER VQQRANCIIG VQTKANCIVG DQRAYGCVLG VQERAGCIVG LQERAGCIIG	LDYPQPIVEH RRYLR RGYPAPIVNH KDYPAPIVNH KDYPRPIVLH KDYPRPIVDH KDYPRPIVDH	R48 N479 N KEARVQTLAA L NLRQKQFKAL AVASKECLAR DSASKECKRK EVVHKENIKR EVVHKENIKR EVVHKKNILR NVVSKQNIQR	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGEAYALNKK MGAAYKVNRE MKAAYAKRSP MKAAYARRSG	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV	472 420 474 549 512 513 497 497
EcCPD TtCPD AnCPD Ds64 At64 Dm64 Dm64 Dm64 X164 AtCRY1	393 361 399 453 416 417 406 406 410	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -YQFNRIY -HQYFRVY -HQYFRVY SREFDRID	NPTTQGEKFD NPASQAKKFD SPVVFGKKYD SPISFGKKYD SPIAFGKKTD SPVAFGKKTD NPQFEGYKFD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRKFL PQGHYIRKYU KHGDYIKKYL PNGEYVRRWL	PELRDVPGKV PEYRSYAPKD PELRHVHPKD PVLKDMPAKY PVLKDMPKQY PELSKYPAGC PVLKKFPTEY PILKKFPAEY PELSRLPTDW	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWTAPKS IYEPWKASLV IYEPWKAPRS IHHPWNAPES	AGVT AR PIER VQQRANCIIG VQTKANCIVG DQRAYGCVLG VQERAGCIVG UQERAGCIIG VLQAAGIELG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH KDYPKPMVLH TDYPHRIVKH KDYPRPIVDH SNYPLPIVGL	R48 N479 N KEARVQTLAA IIIRQKQFKAL AVASKECIAR DSASKECKRK EVVHKENIKR EVVHKKNIIR NVVSKQNIQR DEAKARLHEA	2 483 YEAARKGK ARDLARG MGAAYKATNT MGAAYKATNT MGAAYKATNR MKAAYAKRSP MKAAYARRSG LSQMWQLEAA	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV SRAAI	472 420 474 549 512 513 497 497 502
EcCPD AnCPD Ds64 At64 Dm64 Dr64 X164 AtCRY1 AtCRY1	393 361 399 453 416 417 406 406 410 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -YQFNRIY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPUSFGKKYD SPIAFGKKTD SPVAFGKKTD NPQFEGYKFD NPALQGAKYD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRHFL PQGHYIRKYV KHGDYIKKYL NNGDYIKKYL PNGEYVRRWL PEGEYIRQWL	PELRDVPGKV PEYPSYAPKD PELRHVHPKD PVLKDMPAKY PULKDMPKQY PELSKYPAGC PVLKKFPTEY PILKKFPAEY PELSRLPTDW PELARLPTEW	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASLV IYEPWKASRS IHHPWNAPES IHHPWDAPLT	AGVT ARER VQQRANCIIG VQTKANCIVG DQRAYGCVLG VQERAGCIVG VLQARGCIIG VLQAAGIELG VLQAAGIELG VLKASGVELG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH KDYPRPIVDH RDYPRPIVCH SNYPLPIVGL TNYAKPIVDI	R48 N479 M KEARVQTLAA L NLRQKQFKAL AVASKECIAR DSASKECKRK EVVHKENIKR EVVHKENIKR EVVHKKNIIR NVVSKONIO DEAKARLHEA DTARELLAKA	2 4483 ARDLARG YNQLKAAI MGAAYKATNT MGEAYALNKK MGAAYKVNRE MKAAYAKRSG LSQMWQLEAA ISRTREAQIM	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTV STEGV SRAAI IGAAPDEI	472 420 474 549 512 513 497 497 502 502
EcCPD TtCPD AnCPD Ds64 At64 Dm64 Dm64 X164 AtCRY1 AtCRY2 AtCRY2 AtCRY3	393 361 399 453 416 417 406 406 410 407 486	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD PREDRIP	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPUSFGKKTD SPUAFGKKTD SPUAFGKKTD NPQFEGYKFD NPALQGAKYD SIFKQAQNYD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRKFL POGHYIKKYL KNGDYIKKYL PNGEYVRRWL PEGEYIRQWL PEGEYIRQWL	PELRDVPGKV PEYPSYAPKD PELRHVHPKD PVLKDMPAKY PVLKDMPAKY PELSKYPAGC PVLKKFPAEY PELSRLPFDW PELSRLPFTEW QQLRLPKEK	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWKASLV IYEPWKASLV IYEPWKSPRS IHHPWDAPLS IHHPWDAPLT RHWFGRLMYM	AGVT AR PIER VQTKANCIVG UQTKANCIVG UQERAGCIUG VQERAGCIUG UQERAGCIUG ULQARGIELG DT	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH KDYPKPIVDH KDYPKPIVEH SNYPLPIVGL SNYPLPIVGL TNYAKFIVDI -VVPLKHENG	R48 N479 N KEARVQTLAA L NLRQKQFKAL AVASKECTAR DSASKECTAR DSASKECTAR EVVHKKNITR EVVHKKNITR NVVSKQNIQR DEAKARLHEA DTARELLAKA PMAGGSK	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKNNRE MKAAYAKRSP MKAAYAKRSP MKAAYARRSG LSQMWQLEAA LSQMWQLEAA LSQMWQLEAA SGGGFGGS	GGSAGKASPA MDGKVDEENL VRTCKEEESS EDKTI STEGV SRAA IGAAPDEI 	472 420 474 549 512 513 497 497 502 502 558
ECCPD TtCPD AnCPD Ds64 At64 Dm64 Dm64 Dr64 X164 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 EtCPF1	393 361 399 453 416 417 406 406 410 407 486 437	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD GHELDRID GHELDRIF -RDRYF -RDRYF	NPTTQGEKFD NPVLQGERHD NPASQAKKPD SPUVFGKKYD SPIAFGKKTD SPIAFGKKTD SPIAFGKKTD NPQEGYKFD NPALQGAKYD NIAKQTKDYD NIAKQTKDYD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGKYIRKFL VGGYIKKYL KHGDYIKKYL PNGEYVRWL PEGEYVRWL PEGEYVRWL PAGEYIKTW	PELRDVPGKV PEYPSYAPKD PELRHVHFKD PVLKDMPAKY PULSDMFKQY PELSKYPAGY PELSRLPTDW PELARLPTEW KELAEVPAAY DOGYNMPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWTAPLS IYEPWKASLV IYEPWKASPRS IHEPWNAPES IHHPWDAPLT IHPPWDAPLT IADPNQAPRE	AGVT ARER VQQRANCIIG DQRAYGCULG VQERAGCIUG VLQERAGCIIG VLQAAGIELG VLAAGVELG DT	LDYPQPIVEH RRVLR RGYPAPIVNH RDYPAPIVNH KDYPFNVLH KDYPFNIVH SNYPLPIVGL TNYAKPIVDI -VVPLKHENG LNYPNKLALP	R48 N479 N KEARVQTLAA L NLRQKQFKAL DSASKECTAR DSASKECKR EVVHKKNTLR EVVHKKNTLR EVVHKKNTLR DEAKARLHEA DTARELLAKA PMAGGS RRDFTEMGSP	2 YEAARKGK ARDLARG YNQLKAAT MGAAYKATNT MGAAYKATNT MKAAYAKRSG MKAAYAKRSG LSQMWQLEAA ISRTREAQIM SGGGFRGS -PGPRRGGGG WEANDACH	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI SRAAI IGAPDEI GGR GGR	472 420 474 549 512 513 497 502 502 558 519
EcCPD TtCPD Ds64 At64 Dm64 Dm64 X164 AtCRY1 AtCRY1 AtCRY3 OtCPF1 PtCPF1 DmcPY1	393 361 399 453 416 417 406 406 410 407 486 437 438 430	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD PREDRYF -RINRF -YQYFRCY -PRLDSSLY	NPTTQGEKFD NPVLQGERHD SPVVFGKKYD SPISFGKKYD SPVAFGKKTD SPVAFGKKTD NPQFEGYKFD NPALQGAKYD NIAKQTKDYD SPIAFGKKTD NIAKQTKDYD SPIAFGKKTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGYIRKFL POGYIRKFL NGGYIKKYL PNGEYVRWL PEGEYVAFWL PAGEYIRWU PAGEYIKWU PNGDYIRKWL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKNFPTEY PILKKFPAEY PELSRLPTDW PELSRLPTDW QQLRRLPKEK KELAEVPAAY POFLMNVPKEP	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASLV IYEPWKASLV IYEPWKASPRS IHHPWDAPLT IHPWDAPLT RHWPGRLMYM IADPNQAPRE IYEPWEAPIE	AGVT AR PIER VQQRANCIIG DQRAYGCVLG VQERAGCIVG ULQAAGIELG VLQAAGIELG DT	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVNH TDYPHRIVKH KDYPRPIVDH SNYPLPIVGL TNYAKFIVDI -VVPLKHGNG LNYPNKLALP ENYPHPIVDH	R48 N479 N KEARVQTLAA AVASKECIAR DSASKECKRK EVVHKENTKR EVVHKKNTLR EVVHKKNTLR DTARELLAKA DTARELLAKA DTARELLAKA RRDFTEMGSP KLVSKNNMS	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKAINT MGAAYKAINK MGAAYKVNRE MKAAYARRSG LSQMWQLEAA ISRTREAQIM SGGGFRGS -PGPRRGGGG MKEAIDAGKI	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV SRAAV SRAAI IGAAPDEI GGR	472 420 474 549 512 513 497 502 502 502 558 519 534 528
EcCPD TtCPD AnCPD Ds64 At64 Dm64 Dm64 Dr64 X164 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1 DmCRY1	393 361 399 453 416 417 406 410 407 486 437 438 430 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY -HQFFRVY SREFDRID GHELDRLD PREDRID PREDRYF -QQYFRCY -QQF-FRCY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVVFGKXYD SPISFGKKTD SPVAFGKKTD NPQFEGYKFD NPALQGAKYD SIPKQAQNYD SIPKQAQNYD SPIAFGKKTD CPVAFGERTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGKYIRKFL POGYYIRKYV KHGDYIKKYL PHGEYVRRWL PEGEYVAFWL PAGEYIRWU PNGDYIRKYL PDGTYIRQYU	PELRDVPGKV PEYPSYAPKD PULRDMPAKY PULKDMPAKY PULKKPPTEY PILKKPPTEY PELSRLPTDW PELSRLPTDW QQLRRLPKEK KELAEVPAAY PQFKDMPAKY PELMRVPAEY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASLV IYEPWKASLV IYEPWKASPS IHHPWNAPES IHHPWDAPLT RHWPGRLMYM HADPNQAPRE IYEPWEAPIE VHEPWRMSAE	AGVT AR PIER VQQRANCIIG DQRAYGCVLG VQERAGCIVG LQERAGCIVG VLQAAGIELG VLQAAGIELG VLQAAGIELG DT DRIG LQKKVGVIVG QQEQYECLIG QQEQYECLIG VQKAAKCIIG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH TDYPHRIVKH KDYPRPIVDH SNYPLPIVGL TNYAKPIVDI -VVPLKHGNG ENYPHPIVDH VHYPERIIDL VHYPERIIDL	R48 N479 N KEARVQTLAA JUASKECIAR DSASKECKR EVVHKENIKR EVVHKENIKR EVVHKKNIR NVVSKONIOB DEAKARHEA DTARELLAKA PMAGGSS KLVSKNMSB SMAVKRMLA EAGSRLATER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNR MKAAYARRSG MKAAYARRSG LSQMWQLEAA ISRTREAQIM SGGGFRGS PGPRRGGG MKEAYDAQKN MKSIXOOLSC	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV IGAAPDEI  REPMPANESH PPPHCRFSNE	472 420 474 549 512 513 497 502 558 519 534 528 503
EcCPD TtCPD AnCPD Ds64 At64 Dr64 X164 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1a DrCRY1a DrCRY1a	393 361 399 453 416 407 406 410 407 486 437 438 430 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY -HQYFRVY GHELDRLD GHELDRLD GHELDRLD PREDRYF -RINRF -RINRF -RCFRCY -QQFFRCY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVJFGKKYD SPISFGKKTD SPVAFGKKTD SPVAFGKKTD NPQFEGYKFD NPALQGAKYD SIFKQAQNYD NIAKQTKDYD SPIAFGKKTD CPVALAKRLD CPVSFGRRTD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRKFL KNGDYIKKYL KNGDYIKKYL PNGEYVRRWL PEGEYIRQWL PEGEYIRWL PNGDYIKKWL PDGTYIKQYV PNGDYIRKYL	PELRDVPGKV PEYPSYAPKD PULRDMPAKY PVLKDMPAKY PULKDMPAKY PELSRVPAGC PULKKPPTEY PELSRLPTDW PELSRLPTDW PELARLPTEW QQLRLPKEK KELAEVPAY PELMNVPKEF PVLRGFPAKF	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPLS IYEPWKAPKS IYEPWKAPKS IHEPWNAPES IHHPWDAPLT RHWFGRLMYM IADPNQAPRE VHEPWRMSAE IYDPWNAPES	AGVT AR PIER VQTKANCIVG VQTKANCIVG VQTKANCIVG VQERAGCIVG VLQAAGIELG VLQAKGVELG DTDRIG LQKKVGVIVG QQEQYECLIG VQKVAKCIIG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH KDYPKPNVLH KDYPRPIVDH KDYPKPIVEH SNYPLFIVGL INYPKLANG LNYPNKLALP ENYPHPIVDH VHYPERIIDL VHYPKMVNH	R48 N479 N KEARVQTLAA L NLRQKQFKAL AVASKECIAR DSASKECIAR EVVHKKNILR NVVSKONIQB DEAKARLHEA DTARELLAKA PMAGSS RRDFTEMGSP SMAVKRIMLA AEASRINIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKNNRE MKAAYARNSG MKAAYARNSG MKAAYARKSP MKAAYARKSG -PGPRGGGG -PGPRRGGGG -PGPRRGGGG MKEAYDAQKN MKELRNSLIT MKQIYQQLSC	GGSAGKASPA MDGKVDEENI STEGV STEGV IGAAPDEI 	472 420 474 549 512 513 497 502 558 519 534 528 503 503
ECCPD TtCPD Ds64 At64 Dm64 Dr64 AtCRV1 AtCRV2 AtCRV3 OtCPF1 DmCRV1 DrCRV1a DrCRV1a DrCRV2a	393 361 399 453 416 406 406 406 407 486 437 438 430 407 407	AAPYFRIF AAPYPRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD GHELDRLD FREDRYF -RCDRYF -YQYFRCY -QQFFHCY -QQFFHCY	NPTTQGEKFD NPVLQGERHD SPVVFGKKYD SPISFGKKYD SPIAFGKKTD SPIAFGKKTD NPQEGYKFD NPALQGAKYD NIAKQTKDYD SPIAFGKTD CPVALAKRLD CPVSFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRKWA ATATYIKRWL KEGAYIRKFL POGKYIRKFL POGKYIRKYL KHGDYIKKYL PNGEYVRWL PEGEYVRWL PAGEYIRWL PAGEYIKWV PNGDYIRKYL PNGDYIRRYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKKPFAEY PELSRLPTDW PELSRLPTDW PELARLPTEW KELAEVPAAY PQFKDMPAKY PULRGFPAKY PILRGFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASLV IYEPWKASPKS IHEPWNAPES IHHPWDAPLT IADPNQAPRE IYEPWEAPIE IYEPWKASAE IYDPWNAPES IYDPWNAPES	AGVT ARR VQQRANCIIG UQTKANCIVG UQRAGCIG VQEAAGCIG VLQAAGIELG VLQAAGIELG LQKKVGVIVG QQEQYECLIG VQKAAKCIIG VQAAKCIIG	LDYPQPIVEH RRVLR RGYPAPIVNH RDYPAPIVNH KDYPRPIVH KDYPRPIVGL SNYPLPIVGL TNYAKPIVDI SNYPLPIVGL LNYPNKLALP ENYPHFIVDH VHYPERIDL VHYPMPMVHH VHYPKPMVNH	R48 N479 N KEARVQTLAA OCACONSTREAM NUSKOFKAL DSASKECKR EVVHKENTKR EVVHKENTKR EVVHKENTKR EVVHKENTKR DEAKARLHEA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKANRE MKAAYAKRSG LSQMWQLEAA ISRTREAQIM MKAAYARRSG LSQMWQLEAA ISRTREAQIM MKAIYAGKS -PGPRRGGGG MKEAYDAQKN MKSIRNSLIT MKQIYQQLSC MKQIYQQLSC	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV SRAAI IGAPDEI GGR GGR YRGLGLLASV YRGLGLLASV	472 420 474 512 513 497 502 558 519 534 528 503 503 503 503
EcCPD TtCPD Ds64 At64 Dm64 Dm64 AtCRY1 AtCRY1 AtCRY3 OtCPF1 PtCPF1 DmCRY1a DrCRY1a DrCRY1a DrCRY2b	393 361 399 453 416 406 407 486 437 438 430 407 407 407 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD GHELDRLD PREDRYF -QYFRCY -QQYFHCY -QQFFHCY -QQFFHCY -QQF-FHCY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVVFGKKTD SPIAFGKKTD SPVAFGKKTD NPQFEGYKFD NPALQGAKYD SIFKQAQNYD NIAKQTKDYD SPIAFGKKTD CPVAFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGRYIRKFL POGRYIRKYV KHGDYIKKYL PNGEYVRWL PEGEYVAFWL PAGEYIRWVL PNGDYIRKYL PNGDFIRRYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKMPKQY PULKKPPAEY PELSRLPTDW PELSRLPTDW QQLRRLPKEK KELAEVPAAY PQFKDMPAKY PELMNVPKEF PULRGFPAKY PILRGFPAKY PILRGFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKAPKS IYEPWKASPKS IHHPWNAPES IHHPWDAPLT RHWPGRLMYM IADPNQAPRE IYEPWRAPES IYDPWNAPES IYDPWNAPES IYDPWNAPHD	AGVT AR PIR VQQRANCIIG DQRAYGCVLG VQERAGCIUG VLQAAGIELG VLQAAGIELG VLQAAGIELG DTDRIG LQKKVGVIVG QVGKAKCIIG VQKAAKCIIG VQLAAKCAIG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVNH TDYPHRIVKH KDYPRPIVGL SNYPLPIVGL TNYAKFIVDI -VVPLKHGNG LNYPNKLALP ENYPHPIVDH VHYPERIIDL VHYPMPWVHH VHYPKPMVNH VHYPKPMVNH	R48 N479 N KEARVQTLAA OLASSKECIAR DSASKECKRK EVVHKENIKR EVVHKENIKR EVVHKKNILR DSASKECKRK EVVHKKNILR DTARELIAKA DTARELIAKA DTARELIAKA MAVKRNMLA SMAVKRNMLA AEASRINIER AEASRINIER AEASRINIER AEASRINIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MGAAYKVNRE MKAAYARRSG LSQMWQLEAA ISRTREAQIM MKAIYAQKS MKGIYQQLSC MKQIYQQLSC MKQIYQQLSC	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV SRAAV SRAA GGR GGR GGR GGR YRCLGLLATU YRCLGLLATU YRCLGLLATU	472 420 474 512 513 497 502 558 519 534 528 503 503 503 503 503
EcCPD TtCPD AnCPD Ds64 At64 Dm64 Dm64 St64 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1 DmCRY1B DrCRY2a DrCRY24	393 361 399 453 416 417 406 406 410 407 486 437 438 430 407 407 407 407 409 404	AAPYFRIF AAPYFRVF PKPLKIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRID GHELDRID PREDRYF -YQYFRCY -RLIDSSLVT -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVVFGKXYD SPISFGKKTD SPVAFGKKTD NPALGGAKYD SIPKQAQNYD SIPKQAQNYD SIPKQAQNYD SPIAFGKKTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGYIRKFL POGYIRKYL KNGDYIKKYL PNGEYVRRWL PEGEYVAFWL PAGEYIRWU PNGDYIRKYL PNGDYIRRYL PNGDFIRRYL PNGDFIRRYL	PELRDVPGKV PEYPSYAPKD PULRDMPAKY PULKDMPAKY PULKMPKQY PULKKPPTEY PILKKPPAEY PELSRLPTDW PELSRLPTDW QQLRRLPKEK KELAEVPAAY PQFKDMPAKY PILRGPPAKY PILRGPPAKY PULRDFPAKY PULRDFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASLV IYEPWKASLV IYEPWKAPKS IHHPWNAPES IHHPWDAPLT RHWPGRLMYM HADPNQAPRE IYEPWEAPIE VHEPWRASAE IYDPWNAPES IYDPWNAPDS IYDPWNAPDS IYDPWNAPDS IYDPWNAPDS	AGVT AR PIER VQQRANCIIG DQRAYGCVLG VQERAGCIVG ULQARAGELG VLQAAGELG DT RIG LQKKVGVIVG QQEQYECLIG VQKAAKCIIG VQKAAKCIIG VQLAACCIIG VQLAACCIIG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH TDYPHRIVKH KDYPRPIVDH SNYPLPIVGL TNYAKPIVDI -VVPLKHGNG ENYPHPIVDH VHYPERIIDL VHYPMPWVHH VHYPKPMVNH VHYPKPMVNH VDYPRPIVSH	R48 N479 N KEARVQTLAA JUNE CANNER NUROKOFKAL AVASKECIAR EVUHKENIKR EVUHKENIKR EVUHKKNIKR DSASKECKR EVUHKKNIK DEAKARLHEA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA ALASKINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNR MKAAYARRSG LSOMWQLEAA ISRTREAQIM SGGGFRGS PGPRRGGG MKEAYDAQKN MKSIXQULSC MKQIYQQLSC MKQIYQQLSC MKQIYQQLSR MRQIYQQLSR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV IGAAPDEI 	472 420 474 549 512 502 502 558 502 534 503 503 503 503 503 505 498
EcCPD TtCPD AnCPD Ds64 At64 Dm64 Dr64 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1 DrCRY1a DrCRY1a DrCRY2a DrCRY2a DrCRY2	393 361 399 453 416 406 406 406 437 438 430 407 407 407 407 407 407 409 404 409	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD GHELDRLD GHELDRIF -RCINFF -YQYFRCY -QQFFRCY -QQFFHCY -QQFFHCY -HKYTRIF	NPTTQGEKFD NPVLQGERHD NPASQAKKPD SPUFPGKKYD SPIAFGKKTD SPIAFGKKTD NPALQGAKYD SIPKQAQNYD NIAKQTKDYD CPVALAKRLD CPVSFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLKRWA ATATYIKRWL KEGAYIRKFL PDGKYIRKFL PDGKYIRKYL KNGDYIKKYL KNGDYIKKYL PNGEYIRQWL PEGEYVRWU PNGDYIRKYL PNGDFIRRYL PNGDFIRRYL PNGDFIRRYL PQGEYLRKYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PVLKDMPKQY PULKDPFKYPEY PILKKPPTEY PILKKPPTEY PELARLPTEW QQLRRLPKEK KELAEVPAAY PQFKDMPAKY PILRGPPAKY PILRGPPAKY PVLKNFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKE IYEPWTAPKE IYEPWKASLV IYEPWKAPKS IHIPWNAPES IHHPWDAPLT IHHPWGRLMYM IADPNQAPRE VHEPWRNSAE IYDPWNAPES IYDPWNAPDS IYDPWNAPDS IYDPWNAPDJ IYEPWKAPED IYAPWTAPEE	AGVT AR PIER VQQRANCIVG DQRAYGCVLG VQERAGCIVG ULQERAGCIVG ULQERAGCIVG ULQERAGCIVG ULQEAYECLIG VQEAYECLIG VQLAAKCIIG VQLAAKCIIG VQLAAKCIIG UQLAAGCLIG UQQSQECLIG	LDYPQPIVEH RRYLR RGYPAPIVDH KDYPKPMVLH KDYPKPIVDH KDYPKPIVDH KDYPKPIVDI -VVPLKHONG LNYPNKLALP ENYPHPIVDH VHYPERIIDL VHYPKPMVNH VHYPKPMVNH KDYPPIVSH KDYPLPWVDH	R48 N479 N KEARVQTLAA L NLRQKQFKAL DSASKECKRK EVVHKKNIK EVVHKKNIK UDSASKECKRK EVVHKKNIK DTARELLAKA PMAGSS RRDFTEMGSP KLVSKNNMS KLVSKNNMS AEASRLNIE AEASRLNIE AEASRLNIE AEASRLNIE IEASQRUEL IEASQRUEL	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MGAAYKVNRE MKAAYARKSP MKAAYARKSP MKAAYARKSP MKAAYARKSP MKAAYARKSP MKAAYARKSP MKAAYAAKASP NKAAYAAKASP MKAAYAAKASP MKAYAQLSC MKQIYQQLSC MKQIYQQLSC MKQIYQQLSC MRQVRTEQQT MRUVREQQT	GGSAGKASPA MDGKVDEENL VRTCKEEESS EDKTI STEGV SRAAI IGAAPDEI IGAAPDEI IGAAPDEI GGR GGR YRCIGLLASV YRCIGLLASV YRCIGLLASV YRCIGLLASV YRCIGLLASV TAELTRDV TAQLTTDI	472 420 474 549 512 502 502 558 502 534 503 503 503 505 498 503 505 498 500
EcCPD TtCPD Ds64 AnCPD Ds64 Dr64 X164 AtCRY1 AtCRY2 AtCRY3 OtCPF1 DrCRY1a DrCRY1a DrCRY1a DrCRY2a DrCRY2a DrCRY24 X1CRY1 X1CRY2	393 361 399 453 416 406 410 407 437 438 437 407 407 407 407 407 407 404 404 404	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRLD GHELDRLD PREDRYF -RINRF -YQYFRCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -HHYTRIF -HYY-TRIF -QQFFHCY	NPTTQGEKFD NPJQGERHD SPVJFGKKYD SPJSFGKKYD SPJAFGKKTD SPJAFGKKTD SPJAFGKKTD NPQFEGYKFD NPALQGAKYD NIAKQTKDYD SPJAFGKTD CPVJAFGRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVFGRRTD	HEGEFIRQWL PEGRWLRKWA ATATYIKRWL KEGAYIRKFL POGKYIRKFL POGKYIRKYL KHGDYIKKYL PNGEYVRWL PEGEYVRWL PAGEYIRWL PAGEYIKWV PNGDYIRKYL PNGDFIRRYL PNGDFIRRYL PNGDFIRRYL PNGDFIRRYL PQGEYLRKYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKMPKQY PELSKYPAGC PULKKPPAEY PELSRLPTDW PELSRLPTDW QQLRRLPKEK KELAEVPAAY PQLKGFPAKY PULRGFPAKY PULRGFPAKY PULKNFPAKY PULKAFPSRY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASRS IYEPWKASRS IHEPWNAPES IHEPWNAPES IHEPWNAPES IYEPWEAPIE IYEPWEAPIE IYDPWNAPES IYDPWNAPDS IYDPWNAPDS IYDPWNAPDS IYDPWNAPES IYDPWNAPES IYDPWNAPES IYADFAPEE IYAPWTAPEE IYAPWTAPEE	AGVT ARR VQQRANCIIG VQTKANCIVG DQRAYGCVLG VQRAQGULG VLQAAGIELG VLQAAGIELG VLXASGVELG QQEQYECLIG VQKVAKCIIG VQLAAKCIIG VQLAAKCAIG IQQQSQCLIG VQKAAKCIIG VQKAAKCIIG	LDYPQPIVEH RRVLR RGYPAPIVNH RDYPAPIVNH KDYPRPIVH KDYPRPIVGL TNYAKPIVDI SNYPLPIVGL TNYAKPIVDI UNYPNKLALP ENYPHPIVDH VHYPERIDL VHYPRPMVNH VHYPERWVNH VDYPRPWVNH KDYPLPWVNH IDYPRVVNH	R48 N479 N KEARVQTLAA OSASKECIAE DSASKECKER EVVHKENTKE EVVHKENTKE EVVHKENTKE EVVHKENTKE DTARELIAKA DTARELIAKA DTARELIAKA DTARELIAKA DTARELIAKA AEASRINTEE AEASRINTEE AEASRINTEE AEASRINTEE AEASRINTEE AEASRINTEE AEASRINTEE	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MKAAYAKRSG LSQMWQLEAA ISRTREAQIM MKAAYARRSG LSQMWQLEAA ISRTREAQIM MKAIYARSS PGPRRGGGG MKEAYDAQKN MKQIYQQLSC MKQIYQQLSR MRQIYQQLSH	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV SRAAI IGAPDEI GGR GGR GGR YRGLGLLASV YRGLGLLASV YRGLSLLASV YRGLSLLASV YRGLSLLASV YRGLSLLASV	472 420 474 512 512 5513 497 502 558 509 534 503 503 503 503 503 503 503 503 503 503
EcCPD TtCPD Ds64 At64 Dm64 Dm64 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1a DrCRY1a DrCRY1a DrCRY2a DrCRY2k X1CRY1 X1CRY2	393 361 399 453 416 406 410 407 437 438 437 407 407 407 407 407 407 407 404 406 411 406	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY HQYFRVY SREFDRID GHELDRLD GHELDRLD PREDRYF -QQYFRCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -HKYTRIF -HHYTRIF -HYTRIF PRENRKF	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVVFGKKTD SPIAFGKKTD SPVAFGKKTD NPQFEGYKFD NPALQGAKYD NIAKQTKDYD SPIAFGKKTD CPVAFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVFFGRRTD CPVFFGRRTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGRYIRKFL POGRYIRKYV KHGDYIKKYL PNGEYVRWL PEGEYVRWL PEGEYVAFWL PNGDYIRKYL PNGDYIRRYL PNGDFIRRYL PNGDFIRRYL PSGDYURKYL SGGDYIRKWL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKNPFAC PULKKPPAEY PELSRLPTDW PELSRLPTDW QLIRRLPKEK KELAEVPAAY POFKDMPAKY PULRGPPAKY PILRGPPAKY PULRGPPAKY PULRFPAKY PULKNPFSQY PULKNFPSQY PULKAFPSRY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKAPKS IYEPWKAPKS IHHPWNAPES IHHPWDAPLT RHWPGRLMYM IADPNQAPRE IYEPWKAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES	AGVT AR PIR VQQRANCIIG DQRAYGCVLG VQTKANCIVG DQRAYGCVLG VLQAAGIELG VLQAAGIELG VLQAAGIELG DT	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVNH TDYPHRIVKH KDYPRPIVEH SNYPLPIVGL TNYAKFIVDI -VVPLKHGNG LNYPNKLALP ENYPHPIVDH VHYPRPMVNH VHYPKPMVNH VHYPKPMVNH VHYPKPMVNH KDYPLPVVH KDYPLPVVH	R48 N479 N KEARVQTLAA OLASSKECIAR DSASKECKRK EVVHKENIKR EVVHKENIKR EVVHKENIKR DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA ACASKINIS SMAVKRIMLA AEASRINIER AEASRINIER AEASRINIER IEASQRILEL TTASEHIQU ZEASRMIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MGAAYKVNRE MKAAYARRSG LSQMWQLEAA ISRTREAQIM MKAIYAQLSA KKQIYQQLSC MKQIYQQLSC MKQIYQQLSC MRQUYQLSR MRQUYQQLSR MRQUYQQLSR MRQUYQQLSH MRQUYQQLSH AGGWEKSAR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTV SRAAV SRAAV GGR GGR GGR GGR YRCLGLLAFU YRCLGLLAFU YRCLGLLAFU YRCLGLLAFU TAQLTRDV TAQLTTD1 TAQLTTD1	472 420 474 549 512 5513 497 502 558 509 534 503 503 503 503 503 503 503 503 503 503
EcCPD TtCPD Ds64 At64 Dm64 Dm64 X164 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1B DrCRY18 DrCRY28 DrCRY28 DrCRY28 X1CRY1 X1CRY2 X1CRYD ErCRY18	393 361 399 453 416 407 486 407 486 437 438 430 407 407 407 409 404 406 407 407 407 407 407 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY -HQYFRVY SREFDRID GHELDRID GHELDRLD PREDRYF -YQYFRCY -QQFFRCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVAFGKKTD SPISFGKKTD SPVAFGKKTD NPALGGAKYD SIPKQQQNYD SIPKQQQNYD SIPKQQQNYD SIPKQQQNYD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGYIRKFL POGYIRKYL KNGDYIKKYL PNEEYVRRWL PEGEYVAFWL PAGEYIRWVL PNGDYIRRYL PNGDYIRRYL PNGDFIRRYL PNGDFIRRYL PNGDFIRRYL PSGDYVRRYL SGGDYIRLWV	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKMFPTEY PILKKFPTEY PILKKFPTEY QQLRRLPYEW QQLRRLPYEW QQLRRLPYEW PULKGFPAKY PILRGFPAKY PILRGFPAKY PULKNFPAKY PULKNFPSQY PULKNFPSQY PULKAFPSRY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASLV IYEPWKASLV IYEPWKASRS IHHPWNAPES IHHPWDAPLT RHWPGRLMYM HWPGRLMYM IADPNQAPRE IYDPWNAPES IYDPWNAPES IYDPWNAPES IYEPWKAPED IYEPWKAPED IYEPWKAPEA	AGVT AR PIER VQQRANCIIG DQRAYGCVLG VQERAGCIVG VULQAAGIELG VLLASGVELG DT LQKKVGVIVG QQEQYECLIG VQKAAKCIIG VQKAAKCIIG VQLAAKCAIG IQKQSGCLIG IQKQSGCLIG IQKQSGCLIG IQKAAKCIIG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH TDYPHRIVKH KDYPRPIVDH KDYPRPIVDH INYPKLALP ENYPHPIVDH VHYPERIIDL UNYPKLALP UNYPKYMVNH VHYPKPMVNH VHYPKPMVNH JDYPKPIVSH KDYPFIVSH KDYPFIVSH	R48 N479 N KEARVQTLAA OLASSKECIAR DSASKECKR EVVHKENIKR EVVHKENIKR EVVHKKNIR NVVSKONIOB DEAKARLHEA DTARELLAKA PMAGGSE KLVSKNNMSE SMAVKRMLA AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MGAAYKVNRE MKAAYARRSG LSOMWQLEAA ISRTREAQIM SGGGFRGS PGPRRGGG MKEAYDAQKN MKGIYQQLSC MKQIYQQLSC MKQIYQQLSC MKQIYQQLSC MRQVRTEQQT MRQVREQT	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV IGAAPDEI  REPMPANESH PPPHCRFSNE VRGLGLLAWV YRGLGLLATV YRGLGLLATV YRGLCILASV YRGLCILASV YRGLCILATV	472 420 474 512 512 502 502 502 503 503 503 503 503 503 503 503 503 503
EcCPD TtCPD Ds64 AnCPD Ds64 Dr64 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1 DrCRY1a DrCRY2a DrCRY2a DrCRY2a DrCRY2x X1CRY1 X1CRY2 X1CRY1 ErCRY1a ErCRY1a	393 361 399 416 417 406 410 407 486 437 438 430 407 407 407 407 407 407 407 407 407 40	AAPYFRIF AAPYFRVF PKPLRIF -QYFRVY -HQYFRVY HQYFRVY SREFDRID GHELDRID GHELDRID GHELDRIF -RCINFF -YQYFRCY -QQFFHCY -QQFFHCY -QQFFHCY -HKYTRIF PRENRKF PRENRKF PRENRKF PRENRCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY	NPTTQGEKFD NPAQAKKPD SPVVFGKKYD SPIAFGKKTD SPIAFGKKTD SPIAFGKKTD NPALQGAKYD NIAKQTKDYD CPVGFGRKTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRWA ATATYIRWH PDGKYIRKFL PDGKYIRKFL PDGKYIKKYL KHGDYIKKYL PHGEYVRWL PEGEYVRWL PAGEYIRQWL PHGEYIRQWL PHGDYIRKYL PNGDFIRRYL PNGDFIRRYL PNGDFIRRYL SGGDYIRKYL SGGDYIRKYL PNGDYIRKYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PVLKDMPAKY PULKDMPKQY PELSKYPAGC PULKKPPEY PELSRLPTDW PELARLPTEW QQLRRLPKEK KELAEVPAAY PQFKDMPAKY PULKGPPAKY PVLKAFPSRY PULKAFPSRY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKAPKS IYEPWKASRS IHEPWNAPES IHHPWDAPLT IXPPWKAPKS IYEPWKAPE IYEPWEAPIE IYDPWNAPES IYDPWNAPUS IYDPWNAPDS IYDPWNAPES AHTPWALSNA IYDPWNAPES AHTPWALSNA	AGVT ARR VQQRANCIIG UQTKANCIVG UQRAGCIUG VLQEAGCIUG VLQEAGCIUG VLAAGIELG VLAAGVELG QQEQYECLIG VQKAAKCIIG VQLAAKCIIG VQLAAKCIIG VQLAACCIIG SLAHANLSLG IQKAAKCIIG IQKAAKCIIG	LDYPQPIVEH RRVLR RGYPAPIVNH RDYPAPIVNH KDYPKPWLH TDYPHRIVKH KDYPKPIVEH SNYPLPIVGL TNYAKPIVDI -VVPLKHCNG LNYPNKLALP ENYPHPIVDH VHYPKPWNH VHYPKPWNH VDYPKPWNH KDYPLPWDH ETYPYPIVAH ETYPYPIVAH ETYPYPIVAH	R48 N479 N KEARVQTLAA OCASCA N475 N N475 N N	2 YEAARKGK MGAAYKATNT MGEAYALNKK MGAAYKATNT MGAAYKVNRE MKAAYAKRSG LSQMWQLEAA ISRTREAQIM MKAAYARKSG SGGGFRGS -PGPRGGGG MKEAYDAQKN MKSLRNSLIT MKQIYQQLSC MKQIYQQLSR MRQVRTEQQT MRLVREAQQK MKQIYQQLSR MRQVRTEQQT PAGSWEKSAR MKQIYQQLSH PAGSWEKSAR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI IGAAPDEI IGAAPDEI IGAAPDEI PPHCRFSNE YRGLGLLATV YRGLGLLATV YRGLGLLASV RGK YRGLGLLATV YRGLGLLATV	472 420 474 512 512 502 558 503 503 503 503 505 498 507 499 507 499 503 507
EcCPD TtCPD Ds64 AnCPD Ds64 Dr64 X164 AtCRY1 AtCRY2 AtCRY2 AtCRY3 OtCPF1 DrCRY1a DrCRY1a DrCRY1a DrCRY2a DrCRY24 X1CRY1 X1CRY2 X1CRY1 ErCRY1a ErCRY1b GgCRY1	393 361 399 453 416 417 406 410 407 407 407 407 407 407 407 407 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY HQYFRVY SREFDRID GHELDRLD GHELDRLD PREDRYF -RINRF -YQYFRCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY PRENRKF -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY	NPTTQGEKFD NPVLQGERHD SPVVFGKKYD SPISFGKKYD SPIAFGKKTD SPIAFGKKTD SPIAFGKKTD NPQFEGYKFD NPALQGAKYD NIAKQTKDYD SPIAFGKTD CPVSFGRTD CPVGFGRTD CPVGFGRTD CPVGFGRTD CPVGFGRTD CPVGFGRTD CPVGFGRTD CPVGFGRTD CPVGFGRTD CPVGFGRTD	HEGEFIRQWL PEGRWLRKWA ATATYIKRWL KEGAYIRKFL POGKYIRKFL NGGYYKKYL NGGYYKKYL PAGEYVRWL PEGEYVRWL PAGEYIKWV PNGDYIKKYL PNGDYIRKYL PNGDFIRRYL PNGDFIRRYL PSGDYVRRYL SGGDYIRKYL PNGDYIRKYL PNGDYIRKYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKMPKQY PELSKYPAGC PULKKPPAEY PELSRLPTDW PELARLPTEW QULRLPKEK KELAEVPAAY PULRGFPAKY PULRGFPAKY PULRGFPAKY PULKAFPSRY PULKAFPSRY PULKAFPSRY PULKAFPAKY PULKAFPAKY PULKAFPAKY PULKAFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASRS IHEPWNAPKS IHEPWNAPKS IHEPWNAPKS IHEPWNAPKS IYEPWKAPED IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS IYAPWKAPKA IYAPWKASAE IYAPWKASAE IYAPWKASAE IYAPWKASAE IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS IYDPWNAPKS	AGVT ARR VQQRANCIIG QQRAYGCVLG VQTKANCIVG DQRAYGCVLG VQRAAGIELG VLKASGVELG VLKASGVELG QVLXASGVELG QQEQYECLIG VQKAKCIIG VQLAAKCIIG VQLAAKCAIG IQKQSGCLIG IQKASKCIIG IQKAAKCIIG UQKAAKCIIG UQKAAKCIIG VQLAAKCIIG VQKAAKCIIG VQKAAKCIIG	LDYPQPIVEH RRVLR RGYPAPIVNH RDYPAPIVNH KDYPRPVVH TDYPHRIVKH KDYPRIVKH SNYPLPIVGL TNYAKFIVDI UNYPKHAGN UNYPKHAGN VHYPERIDL VHYPKPWVH VHYPKPWVH VDYPKPWVH KDYPLPWVH KDYPLPWVH KDYPLPWVH KDYPLPWVH VNYPKPWVNH	R48 N479 N KEARVQTLAA OSASKECIAR DSASKECKR EVVHKENTKR EVVHKENTKR EVVHKENTKR EVVHKENTKR EVVHKENTKR DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA RASSRIT DEAKANING AEASRINTER AEASRINTER AEASRINTER PEWSRINOK AEASRINTER DEASRINTER DEASRINTER DEASRINTER	2 YEAARKGK XNQLKAAI MGAAYKATNT MGAAYKATNT MGAAYKATNT MKAAYAKRSG LSQMWQLEAA ISRTREAQIM MKAAYARKSG PGPRRGGGG PGPRRGGGG MKEAYDAQKN MKQIYQQLSC MKQIYQQLSC MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV SRAAI IGAAPDEI GGR GGR GGR YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV	472 420 549 512 502 558 502 558 503 503 503 505 498 500 507 499 503 503 503 503 503 503 503 503 503 503
EcCPD TtCPD Ds64 AnCPD Ds64 At64 Dm64 Dm64 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1a DrCRY1a DrCRY1a DrCRY2b DrCRY24 X1CRY1 X1CRY2 X1CRY1 ErCRY1b ErCRY1b ErCRY1b ErCRY1b GgCRY2 GgCRY2	393 361 399 453 416 417 406 410 407 407 407 407 407 407 409 404 406 4111 410 407 407 407 407 407 407 407 407 407 40	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY HQYFRVY SREFDRID GHELDRLD GHELDRLD PREDRYF -QQYFRCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY PRENRKF -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY	NPTTQGEKFD NPAQAKKFD SPVVFGKKJD SPVVFGKKJD SPIAFGKKTD SPVAFGKKTD NPQFEGYKFD NPALQGAKJD SIPKQAQNYD NIAKQTKDYD SPIAFGKKTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL POGKYIRKFL POGKYIRKYL KHGDYIKKYL PNGEYVRWL PEGEYVRWL PAGEYIRWL PNGDYIRKYL PNGDYIRKYL PNGDFIRRYL PNGDFIRRYL SGGDYIRKYL SGGDYIRKYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKMFPTEY PULKKFPTEY PELSRLPTDW PELSRLPTDW QQLRRLPKEK KELAEVPAAY PQFKDMPAKY PULKGFPAKY PULKGFPAKY PULKNFPSQY PULKNFPSQY PULKNFPSRY PULKNFPSRY PULKGFPAKY PULKGFPAKY PULRGFPAKY PULRGFPAKY PULRGFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKAPKS IYEPWKASRS IHHPWNAPES IHHPWDAPLT RHWFGRLMYM IADPNQAPRE IYEPWRAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES	AGVT ARR VQQRANCIIG DQRAYGCVLG VQTKANCIVG DQRAYGCVLG VLQAAGIELG VLQAAGIELG VLQAAGIELG DT	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVNH TDYPHRIVKH KDYPRPIVEH SNYPLPIVGL TNYAKFIVEH SNYPLPIVGL TNYAKFIVDI UNYPNKLALP ENYPHPIVDH VHYPRPMVNH VHYPKPMVNH VDYPKPMVNH KDYPPIVSH KDYPPIVSH KDYPPIVSH KDYPPIVSH VNYPKPMVNH VNYPKPMVNH VNYPKPMVNH VNYPKPMVNH	R48 N479 N KEARVQTLAA JCASKECIAK DSASKECKRK EVVHKENIKR EVVHKENIKR EVVHKENIKR DSASKECKRK EVVHKENIKR INVSKQIIOE DEAKARLHEA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA AEASRINIER AEASRINIER AEASRINIER JEASGRUIEL TTASEHTULI TASENTURE PEWSRHINGK AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MGAAYKATNRSG LSQMWQLEAA ISRTREAQIM MKAAYARRSG LSQMWQLEAA ISRTREAQIM MKAIYQLSA MKQIYQQLSC MKQIYQQLSC MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTV SRAAV SRAAV GGR GGR GGR GGR	472 420 474 549 512 552 552 552 553 503 503 503 503 503 503 503 503 503
EcCPD TtCPD Ds64 At64 Dm64 Dr64 AtCRY1 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1a DrCRY1a DrCRY1a DrCRY2a DrCRY24 DrCRY2 X1CRY1 X1CRY2 X1CRY1 ErCRY1a ErCRY1a ErCRY1a	393 361 399 453 416 406 407 406 407 407 407 407 407 407 407 407 407 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY HQYFRVY SREFDRID GHELDRLD PREDRIP PREDRYF -QQYFRCY -QQFFHCY	NPTTQGEKFD NPVLQGERHD NPASQAKKFD SPVAFGKKTD SPIAFGKKTD SPVAFGKKTD NPALQGAKYD SIPKQAQNYD SIPKQAQNYD SIPKQAQNYD SIPKQAQNYD SPIAFGKKTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRRWA ATATYIKRWL KEGAYIRKFL PQGHYIRKYV KHGDYIKKYL PNGEYVRRWL PEGEYVAFWL PAGEYIRWVL PNGDYIRRYL PNGDYIRRYL PNGDFIRRYL PNGDYIRRYL PSGDYVRRYL PSGDYVRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKKFPTEY PILKKFPTEY PELSRLPTDW PELARLPTEW QQLRRLPKEK KELAEVPAAY PULRGFPAKY PULRGFPAKY PULRGFPAKY PULKNFPSQY PULKNFPSQY PULRGFPAKY PILRGFPAKY PILRGFPAKY PILRGFPAKY PILRGFPAKY PILRGFPAKY PILRGFPAKY PILRGFPAKY PILRGFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPLS IYEPWKAPLS IYEPWKAPKS IHHPWNAPES IHHPWDAPLT RHWPGRLMYM RHWPGRLMYM RHWPGRLMYM IYDFWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDFWNAPES IYDFWNAPES	AGVT AR PIER VQQRANCIIG DQRAYGCVLG VQTKANCIVG DQRAYGCVLG VULDAAGIELG VLLASGVELG DT DRIG LQKKVGVIVG QQEQYECLIG VQKAAKCIIG VQKAAKCIIG IQKQSGCLIG IQKAAKCIIG IQKAAKCVIG VQKAAKCVIG VQKAAKCVIG VQKAAKCVIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG	LDYPQPIVEH RRYLR RGYPAPIVNH RDYPAPIVDH TDYPHRIVKH KDYPRPIVDH KDYPRPIVDH KDYPRPIVDH INYPKKLALP ENYPHPIVDH VHYPKPWVNH VHYPKPWVNH VHYPKPWVNH IDYPRIVSH KDYPFIVSH KDYPFIVSH VNYPKPWVNH VNYPKPWVNH VNYPKPWVNH VNYPKPWVNH VDYPFPWVDH	R48 N479 N KEARVQTLAA JUASKECIAR DSASKECKR EVVHKENIKR EVVHKENIKR EVVHKKNIR NVVSKONIOB DEAKARLHEA DTARELLAKA PMAGGSE KLVSKNNMSE SMAVKRMLA AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MGAAYKVNRE MKAAYARRSG LSOMWQLEAA ISRTREAQIM SGGGFRGS PGPRRGGG MKEAYDAQKN MKGIYQQLSC MKQIYQQLSC MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV IGAAPDEI GGR REPMPANESH PPPHCRPSNE GGRLAMV YRGLGLLAMV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV	472 420 474 512 512 502 502 502 503 503 503 503 503 503 503 503 503 503
EcCPD TtCPD Ds64 AnCPD Ds64 Dr64 AtCRY1 AtCRY1 AtCRY2 AtCRY3 OtCPF1 PtCPF1 DmCRY1a DrCRY2a DrCRY2a DrCRY2a DrCRY2x X1CRY1 X1CRY2 X1CRY1 X1CRY2 X1CRY1 S1CRY2 GGCRY1 GGCRY1 GGCRY2 GGCRY4 MmCRY2	393 399 453 416 407 407 407 407 407 407 407 407 407 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY HQYFRVY SREFDRID GHELDRLD GHELDRLD GHELDRLD -RCINRF -YQYFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -HKYTRIF -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY -QQFFRCY	NPTTQGEKFD NPLQGERHD NPASQAKKPD SPUFFGKKTD SPIAFGKKTD SPIAFGKKTD SPIAFGKKTD NPALQGAKTD NPALQGAKTD CPVAFGKKTD CPVAFGRKTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRWA ATATYIRWL KEGAYIRKFL POGYIRKFL POGYIRKYL KNGDYIKKYL PNGFYRWL PAGEYIRWU PAGEYIRWU PNGDYIRKYL PNGDFIRRYL PNGDFIRRYL PNGDFIRRYL PSGDYURKYL PNGDYIRKYL PNGDYIRKYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL PNGDYIRRYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKDMPKQY PELSKYPAGC PULKKPPEY PELSRLPTDW PELARLPTEW QQLRRLPKEK KELAEVPAAY PQFKDMPAKY PULKGPPAKY PULKDPPAKY PULKAPPSRY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGPPAKY PULKGFPSRY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKAPKS IYEPWKASRS IHHPWNAPES IHHPWDAPLT IYEPWKASRS IHHPWDAPLT IYEPWEAPIE IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES	AGVT ARR VQQRANCIIG VQTKANCIVG UQRAGCIIG VLQAAGIELG VLQAAGIELG VLKASGVELG QQEQYECLIG VQKAAKCIIG VQLAAKCIIG VQLAAKCIIG VQLAAKCIIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG IQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG	LDYPQPIVEH RRVLR RGYPAPIVNH RDYPAPIVNH KDYPRPIVH KDYPRPIVEH SNYPLPIVGL -VVPLKHGNG LNYPNKLALP ENYPHPIVDH VHYPERIDL VHYPENVNH VDYPRPVNH KDYPLPWVNH IDYPRPIVSH ETYPYPIVNA ETYPYPIVNH VNYPKPWVNH VNYPKPWVNH VNYPKPWVNH VNYPKPWVNH VNYPKPWVNH VNYPKPWVNH VNYPKPWVNH	R48 N479 N KEARVQTLAA OLASSKECIAR DSASKECIAR DSASKECKR EVVHKKNIKR EVVHKKNIKR DEAKARLHEA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DEASRINIER AEASRINIER AEASRINIER AEASRINIER DEWSRINICK AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER AEASRINIER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGAAYKATNT MKAAYAKRSG LSQMWQLEAA ISRTREAQIM MKAAYARRSG LSQMWQLEAA ISRTREAQIM MKAIYAQLSC MKQIYQQLSC MKQIYQQLSR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI IGAAPDEI IGAAPDEI IGAAPDEI GGR GGR GGR YRGLGLLASV YRGLGLLASV YRGLGLLASV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLATV YRGLGLLASV IAQLTRDD YRGLGLLASV	472 420 474 549 513 497 502 558 503 503 503 505 505 505 498 500 505 505 505 505 505 505 505 505 50
EcCPD TtCPD Ds64 AnCPD Ds64 Dr64 X164 AtCRY1 AtCRY2 AtCRY2 AtCRY3 OtCPF1 DrCRY1a DrCRY1a DrCRY1a DrCRY2a DrCRY2a DrCRY2 X1CRY1 X1CRY2 X1CRY1 ErCRY1a ErCRY1b GgCRY1 GgCRY1 GgCRY1 MmCRY2 HsCRY1	393 399 453 416 407 406 407 408 407	AAPYFRIF AAPYFRVF PKPLRIF -AQYFRVY -HQYFRVY HQYFRVY SREFDRID GHELDRLD GHELDRLD PREDRYF -RINRF -YQYFRCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY -QQFFHCY	NPTTQGEKFD NPVLQGERHD SPVVFGKKTD SPISFGKKTD SPIAFGKKTD SPIAFGKKTD SPIAFGKKTD NPQFEGYKFD NPALQGAKYD NIAKQTKDYD SPIAFGKTD CPVSFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD CPVGFGRRTD	HEGEFIRQWL PEGRWLRKWA ATATYIRKWL KEGAYIRKFL POGYIRKFL POGYIRKYL KHGDYIKKYL PNGEYVRWL PEGEYVRWL PAGEYIKWV PNGDYIRKYL PNGDYIRKYL PNGDFIRRYL PNGDFIRRYL PSGDYVRYL SGGDYIRKYL PNGDYIRRYL PNGDYIRRYL	PELRDVPGKV PEYPSYAPKD PULKDMPAKY PULKDMPAKY PULKDMPAKY PELSRVPAGC PULKKPPAEY PELSRLPTDW PELARLPTEW QULRLPKEK KELAEVPAAY PULRGFPAKY PULRGFPAKY PULKDFPAKY PULKAFPAKY PULKGFPAKY PULKGFPAKY PULKGFPAKY PULKGFPAKY PULKGFPAKY PULKGFPAKY	VHEPWKWAQK PVVDLEE LISGEIT IYEPWTAPKS IYEPWKASRS IHEPWNAPES IHEPWNAPES IHEWNAPES IYEPWKASRS IYOPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYDPWNAPES IYEPWNASEE IYDPWNAPES IYEPWNASE	AGVT ARR VQQRANCIIG VQTKANCIVG DQRAYGCVLG VQRAQGULG VLKASGVELG VLQAAGIELG VLVAAGIELG VLVAAGUELG QQEQYECLIG QQEQYECLIG VQKAAKCIIG VQLAAKCAIG VQLAAKCAIG VQLAAKCIIG IQKAAKCIIG UQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG UQKAAKCIIG UQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG VQKAAKCIIG	LDYPQPIVEH RRVLR RGYPAPIVNH RDYPAPIVNH KDYPRPVVH TDYPHRIVKH KDYPRPIVEH SNYPLPIVGL TNYAKFIVDI UNYPKLAUP ENYPHPIVGL VHYPERIDL VHYPKHVNH VHYPKPWVNH VDYPKPWVNH KDYPLPWVH VNYPKPWVNH VNYPKPWVNH VDYPKPWVNH VDYPKPWVNH VDYPKPWVNH VDYPKPWVNH VDYPKPWVNH VDYPKPWVNH VDYPKPWVNH VDYPKPWVNH	R48 N479 N KEARVQTLAA OCASCA N475 N KEARVQTLAA AVASKECIAE DSASKECKR EVVHKENTKR EVVHKENTKR EVVHKENTKR EVVHKENTKR DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA DTARELLAKA NUVSKONTOR DEAKARLEA REASRINTER AEASRINTER AEASRINTER AEASRINTER AEASRINTER AEASRINTER AEASRINTER AEASRINTER AEASRINTER AEASRINTER	2 YEAARKGK ARDLARG YNQLKAAI MGAAYKATNT MGEAYALNKK MGAAYKVNRE MKAAYARSG LSQMWQLEAA ISRTREAQIM MKAAYARSG LSQMWQLEAA ISRTREAQIM MKAAYARSG PGPRRGGGG MKEAYDAQKN MKQIYQQLSC MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR MKQIYQQLSR	GGSAGKASPA MDGKVDEENL VRTGKEEESS EDKTI STEGV SRAAI IGAAPDEI GGR GGR GGR GGR YRGLGLLASV YRGLGLLASV YRGLGLLASV YRGLGLLASV YRGLGLLASV YRGLGLLASV YRGLGLLASV	472 420 474 549 512 5513 497 502 558 519 534 503 503 503 505 505 505 505 505 505 505

**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 4<sup>th</sup> 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 4<sup>th</sup> 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	XI = Xenopus laevis	Ot = Ostreococcus tauri	Pt = Phaeodactylum tricornutum		
Er = Erithacus rubecula	Gg = Gallus gallus	Mm = Mus musculus	Hs = Homo sapiens			
CPD = Cyclobutane Pyrimidi	ne Dimer photolyase	64 = (6-4) photolyase				
CRY = Cryptochrome		CRYD = "DASH" cryptochrome				
CPF = Cryptochrome Photoly	yase Family					

# **Supplemental Transient Absorption Data**



**Figure S3.** Nano- and microsecond transient absorption kinetics. Flash-induced absorption changes were recorded for (a) WT X/(6-4)PL and (b) its W370F mutant at the indicated wavelengths. FAD<sub>ox</sub> was excited by a laser pulse at t = 0 (355 nm, 100 ps pulse duration,  $E \sim 4.0$  mJ.cm<sup>-2</sup>. 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.

Protoin	Timescale	۶.	$ au_2$	
Frotein	(Fitted data)	21		
)A/T	ns/µs	$(2.556 \pm 0.000) \times 10^{-6}$ c		
W I	(Figure S3a)	(2.330 ± 0.009) × 10 5	-	
\ <b>\</b> /T	ms/s	$(3.423 \pm 0.005) \times 10^{-2}$ s	_	
W I	(Figure S4a)	(3.423 ± 0.003) × 10 3	-	
W/270E	ns/µs	$(1.996 \pm 0.015) \times 10^{-7}$ c	(1.052 ± 0.003) × 10 <sup>-5</sup> s	
W3/UF	(Figure S3b)	(1.000 ± 0.013) × 10 5		

**Table S1.** Fit results. Time constants  $\tau_1$  and  $\tau_2$  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 *XI*(6-4)PL (20  $\mu$ M 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<sub>ox</sub> was excited by a laser pulse at *t* = 0 (355 nm, 100 ps duration, *E* ~4.0 mJ.cm<sup>-2</sup>). Global fit of the decay curves yields a time constant  $\tau$  ~35 ms (see Table S1). Inset: Expansion of the 590 nm trace showing an absorption increase with  $\tau$  ~35 ms attributed to formation of the neutral FADH<sup>•</sup> 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*<sub>4</sub> *radicals and the deprotonation kinetics of Trp*<sub>4</sub>*H*<sup>\*</sup>. According to the published spectra of (free) TrpH<sup>\*+</sup> and Trp<sup>\*</sup>, deprotonation of TrpH<sup>\*+</sup> 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<sup>--</sup>TrpH<sup>\*+</sup> and the FAD<sup>--</sup>Trp<sup>\*</sup> pair (Figure 4). These observations indicate that the environment of the 4<sup>th</sup> 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<sup>-</sup> groups of D318 and D320, C(NH<sub>2</sub>)=NH<sub>2</sub><sup>+</sup> 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<sup>\*+</sup> and Trp<sup>\*</sup> around 450 nm (*cf.*, methionine to Cu<sup>2+</sup> CT transitions in plastocyanin and azurin)<sup>7</sup>.

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

Trp<sub>3</sub>H<sup>\*+</sup> in the W370F mutant protein (see main text), in *E.coli* PL<sup>8</sup> and in *At*Cry1<sup>3a</sup>). A similarly slow TrpH<sup>\*+</sup> deprotonation (2.56 µs) has been reported very recently for *Drosophila melanogaster* cryptochrome.<sup>9</sup> Interestingly, *Dm*Cry also possesses the 4<sup>th</sup> Trp (W394 in *Dm*Cry; see Figures 2 and S2). We consider it very likely that W394 (and not the 3<sup>rd</sup> Trp W342 as assumed in Ref. <sup>9</sup>) functions as the terminal member of the Trp chain in *Dm*Cry. In line with our model, the FAD<sup>+-</sup>Trp<sup>+</sup> radical pair observed after deprotonation of TrpH<sup>++</sup> recombined rather slowly in *Dm*CRY (6.8 ms at pH 7.0).<sup>9</sup> In bird (garden warbler) CRY1, the FAD<sup>+-</sup>Trp<sup>+</sup> pair was reported to recombine in 14 ms (at pH 7.4).<sup>10</sup> 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.<sup>8, 11</sup> 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*<sup>--</sup> *Trp*<sup>-</sup> and *FAD*<sup>--</sup> *protonation*. The signals on the 0.4 s timescale (Figure S4a) are dominated by decay consistent with recombination of an FAD<sup>--</sup> Trp<sup>-</sup> 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<sup>-</sup> Trp<sup>-</sup> pair. The FADH<sup>-</sup> absorption at 590 nm (inset in Figure S4a) built up with the same kinetics as the decay of FAD<sup>--</sup> Trp<sup>-</sup>. Within a simplified parallel reaction scheme, the initial to final amplitude ratio of 6 to 1 at 450 nm (note that FAD<sup>--</sup> and FADH<sup>-</sup> absorb about equally at 450 nm (see Figure 4a) yields intrinsic time constants of ~40 ms for recombination of FAD<sup>--</sup> Trp<sup>-</sup> and ~200 ms for protonation of FAD<sup>--</sup>. Noteworthy, protonation of FAD<sup>--</sup> 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 *XI*(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 cofactor to FAD<sup>--</sup> in plant Crys.<sup>3a, 12</sup>

# References

- 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.
- (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.