Supplementary information

Supplementary Materials and Methods

Phylogenetic analysis of class II photolyases

Protein sequences of the whole photolyase-cryptochrome family were derived from the UniProtKB database (UniProtConsortium, 2009) and aligned using ClustalW2 (Larkin et al, 2007). A phylogenetic tree was constructed from multiple alignment output data using PHYLIP tree type with Kimura correction of distances, no ignoring of gaps in alignment and clustering set to neighbor joining (Saitou & Nei, 1987). Visualization of the unrooted tree was done with MEGA4 software (Tamura et al, 2007). Sequence identities were taken from ClustalW2 output. The analysis of the class II subfamily utilized a set of 98 non-redundant sequences harbored by the class II specific INTERPRO profile 008148 (Hunter et al, 2009), subjected them to a multiple sequence alignment by KALIGN (Lassmann & Sonnhammer, 2005), removal of gaps and final phylogenetic analysis with ClustalW2.

Generation and preparation of MmCPDII mutants

Photolyase mutants *Mm*CPDII-W388F, *Mm*CPDII-W360F, *Mm*CPDII-W381F, *Mm*CPDII-Y380F and *Mm*CPDII-Y345F as well as *Mm*CPDII-N403D, *Mm*CPDII-N403L and *Mm*CPDII-N403A were obtained from pET-28a-*Mm*CPDII by site-directed mutagenesis using Phusion[®] DNA polymerase (Finnzymes) and primer pairs as listed in Suppl. Table SII. Subsequently, resulting plasmids were verified by sequencing (Qiagen). Gene expression of photolyase mutants was performed analogously to wildtype, but soluble protein was produced at different temperatures (for details see Suppl. Table SII). All mutants could be purified by NiNTA affinity and size exclusion chromatography according to the procedures described above. The W381F mutant was especially prone to aggregation and had to be characterized shortly after purification.

UV/Vis spectra, photoreduction and CPD repair activity of MmCPDII

Absorption spectra were recorded using a DH-2000-BAL light source (Mikropack) and Maya 2000 Pro spectrometer (Ocean Optics) and a V-660 spectrometer (JASCO), respectively. For photoreduction, the protein solved in buffer III (10 mM Tris-HCl, 100 mM NaCl, pH 8.0, 25 mM DTT) was incubated for 5 min in the dark and the first spectrum at t = 0 min was taken afterwards. Subsequently, the sample was continuously illuminated with a high power LED at 450 nm (9.7 mW cm⁻² at 10 cm, Roithner Lasertechnik) and spectra were recorded as mentioned in the illustrations. To determine half-times of photolyase wildtype and mutants, absorbance at 450 nm during photoreduction was recorded in intervals of one second for three times per sample. Subsequently, the data was fitted by a first order exponential decay. A modified DNA-repair assay (Jorns et al, 1985) was performed with the *M. mazei* photolyase and the CPD-lesion containing oligo(dT)₁₈ in buffer III containing 4.6 µM protein and 5 µM CPD-lesion (final concentrations). After 5 min of dark incubation, the assay was illuminated with a high power LED (Roithner Lasertechnik) at 395 nm (0.9 mW cm⁻² at 10 cm) and repair activity was observed at 265 nm.

Electrophoretic mobility shift assay (EMSA) of MmCPDII•CPD-DNA complexes

Manipulations with IRDye700-labeled DNA probe alone were performed under green light to prevent degradation of the fluorophore whereas reactions containing also the protein were handled under red light to prevent any repair event. Synthesis of the 50mer oligonucleotide with a single

5'-IRDye700-AAAATGCTGGATGTCGAGGTGTAAT<>TAATGTGGAGCTGTAGGTCGTA AAA-3', was done according to the published procedure (Pokorny et al, 2008). For EMSA, the **CPD**-comprising oligonucleotide was annealed with the complementary strand, 5'-TTTTACGACCTACAGCTCCACATTAATTACACCTCGACATCCAGCATTT-3'. The binding reaction was carried out on ice in buffer IV (10 mM Tris-HCl, 100 mM NaCl, pH 8.0, 10% glycerol) for 30 min in the dark. The assays with a total volume of 10 µl contained 2 nM duplex DNA and increasing photolyase concentrations (0, 5, 10, 50, 100, 200, 300, 400, 500, 600, 800 and 1000 nM, respectively). Binding reactions were analyzed by a native PAGE (5% polyacrylamide) in TBE buffer (25 mM Tris, 25 mM boric acid, 0.625 mM EDTA, pH 8.4) at 4 °C and resulted bands were quantified by the Odyssey Imaging System (LI-COR Biosciences). Binding assays were repeated three times and data were fit with Origin 7.0

(Microcal) using the simple Hill-equation (Hill, 1910), $\theta = \frac{[DNA]^n}{K_{D,NS}^n + [DNA]^n}$, or a mixed one with

specific binding, $\theta = 0.5 \bullet \left(\frac{[DNA]}{K_{D,S} + [DNA]} + \frac{[DNA]^n}{K_{D,NS}^n + [DNA]^n} \right)$, to determine dissociation

constants.

Supplementary Figures



Figure S1. *M. mazei* class II photolyase complexed to CPD-DNA. Electron density (SIGMAA-weighted $2F_{obs}$ - F_{calc}) is contoured at 1 σ . (A) The 14mer duplex DNA comprising the chemically synthesized CPD-lesion is bound to the catalytic domain of the photolyase. The thymine dimer (T7<>T8) is flipped into the binding pocket and is located close to the adenine moiety of the catalytic cofactor FAD. Colouring corresponds to the overall structure of the uncomplexed photolyase. (B) Complexation of *Mm*CPDII and CPD-DNA goes along with a

quasi-continuously arrangement of the dsDNA in the asymmetric unit. (C) Duplex CPD-DNA taken from crystal structures of *Mm*CPDII•CPD-DNA and *An*CPDI•CPD-DNA were superimposed onto the CPD-lesions and elongated at their 5'- and 3'-arms with modeled duplex B-DNA afterwards. The overall bend of duplex B-DNA with an internal CPD lesion bound to *Mm*CPDII (green) differs from *An*CPDI. The CPD-DNA is kinked from about 27° - 30° in the unbound state (Husain et al, 1988; Park et al, 2002; Pearlman et al, 1985) to about 50° by flipping the thymine dimer into the active site (Mees et al, 2004). CPD-DNA bound to the *M. mazei* class II photolyase is additionally dislocated by 45° relative to the *A. nidulans* complex. (D) In contrast to other photolyase subclasses, *Mm*CPDII lacks a C-terminal helical extension, whereas the Western rim is built by the loop between helices $\alpha 17$ and $\alpha 18$. (E) Structural comparison of *Mm*CPDII (pale green) to the *Mm*CPDII•CPD-DNA complex (green) shows conformational changes of the "bolt-like" amino acids D428, R429, W431 and R441.



Figure S2. *Mm*CPDII crystals and *in crystallo* UV/Vis spectroscopy. (A and B) Crystals of *M. mazei* class II photolyase and (C and D) in complex with CPD-DNA. Images of crystals

mounted in cryo-loops were taken at ESRF (Grenoble, France) prior to data collection. Crystals of *Mm*CPDII•CPD-DNA were documented and mounted in a cryo-loop under red-light to avoid light-induced repair. (E) *Mm*CPDII•CPD-DNA crystals were cryosoaked in crystallization buffer supplemented with 30% glycerol and 10 mM DTT and flash-frozen in liquid nitrogen. The first spectrum (black) was recorded before any illumination and exhibits characteristic absorption peaks for *Mm*CPDII in the oxidized state. Illumination at 450 nm under cryogenic conditions (blue) shows no effect. Photoreduction via illuminated at room temperature. (F) X-ray radiation (flux: ~5.0•10¹⁰ photons/s) for 300 seconds on *Mm*CPDII•CPD-DNA crystals flash-frozen without supplemented DTT causes a loss of absorbance between 400 nm and 500 nm indicating a reduction of the catalytic cofactor FAD. All *in crystallo* UV/Vis spectra were recorded at 100 K using a DH-2000-BAL light source (Mikropack) and a HR2000 spectrometer (Ocean Optics) at beamline ID14-1 (ESRF, Grenoble, France). Spectra were smoothed three times by adjacent averaging of three data points.



Figure S3. Angular analysis of CPD-lesions. Given angles between the 5'- and 3'-thymidine base planes were calculated using two normal straight lines (2XRZ: *Mm*CPDII•CPD-DNA; CCDC100804: Synthetic CPD-lesion; 3MR3, 3MR4: Human DNA polymerase η•CPD-DNA; 1TEZ: *An*CPDI•CPD-DNA; 2VTB: *At*DASH•CPD-DNA).

Α

MacPDII 217 LLINNRDLFE FWHF-EFGE KAARKYMEOF IADRLDDY GALEDTTKMURADNLS TYLHFQQIDS QNVULEVEKA EDN	
ECCEDI 166 -TILMYPROS FOTAMFPVEE KAAIAQUROF CONGAGEY EQGREFPAY -TESTERIS ASLATGGLEP ROC.HALLAE QPQALOG GAGSWUNNE 274 AnCEDI 200 DOG GFPVEFGE TAAIALUGF CORAIATADY PORNFAREAGTGGLES PALKEGAIG LOGAWQASASA HALGEN-GENE NSIEWWOOL 283 FCODI 172 DEPLEPERGE TAAIALUGF CORAIATADY PORNFAREAGTGGLES PALKEGAIG LOGAWQASASA HALGEN-GENE NSIEWWOOL 283 ACDASH 243	
AncPDI 200 WDG GFPVERGE TAALARLOFF CORALAPY DPQNNFFAE- GEGGREGAL REQARGALEA HALER-SOEAR NEIEWWOOF 283 TCCPDI 172	
TrCDDI 172	
ACDASH 243TR GWRFVGGE SAGUGRVFEY FWKKULLKVY KETNOMUGPDYSTKFS FWLAFGCISP RFIFEDVORY EKERVAN NSTYWLFF 25 Dm(6-4) 215	
ALDAH 213 IN GWAT VOSE JANGKANINI MIKULANINI KAINANANA PIDIALS HURKOLDA NI HELENANA KAINANA 229 MmCPDI 302 ILIWKEISDN FCYTNPGYDGFESFFSW AKESIAAHEN DYRSHITLE EFEAGKTHDJ INASOMELL STGKMMENTE MIKASKILE SESPEKALE 397 ECCPDI 275 LI-WREFYCH ALTYNPSICKHERFIANT DRV GWOSNIPLE EFEAGKTHDJ INASOMELL STGKMMENER MITASFLV-K DLLIDER 360 ACCPDI 244 LA-WREFYCH ALTYNPSICKHERFIANT DRV GWOSNIPLE AFEAGACHDO AWGEGKTGYD IVDAAMKELL STGKMMENER MITASFLV-K DLLIDER 360 TCCEDI 244 LA-WREFYCH ALTYNPENCKHERFIANT DRV GWOSNIPLEDA AWGEGKTGYD IVDAAMKELL AGTELSTRAR MINASFLV-K DLLIDER 360 ACDAH 326 LI-WRDYFRF LSIKCGKSLHILGGFRAW Q GWOSDAKES SUDAKTGYD LIDAAMKELS TEGMSMERGE QUVGSEUV-R DMGLORE 410 Dm(6-4) 300 LM-WREFYYT VAAAEPNFOR MLGRVYCMQI P QEHEDHE AWTHGRTGYD FIDALMRCLR QEGWIHLAR HAVACFLTRG DLWISNE 385	
Dam(0-4) 213	
302 ILIWKEISON FCYYNPGYDGFESFFSW AKESINAHEN DYRSHITTLE EFEAGKTHDP INNASOMELL STGKMHEYTR MUAKKIL-E SESPEKALE 397 ECCEDI 275 LI-WREFYRH LITYHPSLOKHRPFIAMT DRV OWOSNPAHLQ AMQEGKTGYP IVDAAMRQLN STGKMHMKIR MITASFLV-K DLLIDWR 360 AnCEDI 284 LA-WREFYRH LITYHPSLOCHRRPFIAMT DRV OWOSNPAHLQ AMQEGKTGYP IVDAAMRQLT ETGKMHMKIR MITASFLV-K DLLIDWR 360 TCCEDI 284 LA-WREFYRH LITYHPSLOCHRRPFIAMT DRV OWOSNPAHLQ AMQEGKTGYP IVDAAMRLH ATGFLSTRAR MITASFLV-K DLLIDWR 368 TCCEDI 284 LL-WREFYRH LITYHPRACHRLDRF RATELPR AMYERLED AMYERKTGYP LIDAMMKLH ATGFLSTRAR MITASFLV-K DLLIDWR 369 ACDASH 325 LI-WRDYFRF LSIKCGKSLHILGGPRWV QG KNSQDQKLFE SKRDAKTGYP LIDAMMKELS TTGFMSNRGR QIVCSELV-R DMGLDRF 410 Dm(6-4) 300 LM-WREFYT VAAAREPNFDR MLGNVYCMQI P QEHEDHLE AWTHGRTGYP FIDAIMRCLR QEGWIHLAR HAVACFLTRG DLWISKE 385	
MmcPDII 302 ILIWKEISDN FCYYNPGYDGFESFFSW AKESINAHRN DYRSHITLE EFEAGKTHDP I NASOMELL STGKMHEYTR MTHAKKIL-E SESPEKALE 397 EGCPDI 25 LI-WREFYRH LITYHPELCKHRPFIAMT DRV GUOSNPAHLQ ANGCORGYD I UDAANROLN STGKMHENLR MITASEIU-K DLLIDAR 360 AGCPDI 284 LI-WREFYRH LITYHPELCGPYRSIMQ QFP GUOSNPAHLQ ANGCORGYD I UDAANROLN STGKMHENLR MITASEIU-K DLLIDAR 368 TCCPDI 245 LL-WRDFSYH LUHFPMAAGPYRSIMQ QFP GUOSNPAHLQ ANGCORGYD I UDAANROLN TEGMMENCR MINASEIT-K DLLIDAR 368 TCCPDI 245 LL-WRDFSYH LUHFPMAAGPYRSIMQ QFP GUOSNPAHLQ ANGCATGYP IUDAANROLN TEGMMENCR MINASPAYKH LLLIDAR 329 AtDASH 326 LI-WRDFSYH LUHFPMAAGRENDR GAUSANG QUUSSELLQ ANGESTGYP LIDAANRELS TIGFMSNRGR QUUSSELV-R DMGLDAR 410 Dm(6-4) 300 LM-WREFYYT VAAAEPNFDR MLGNVYCMQI P "QEHPDHLE AWTHGRTGYP FIDAIMRQLR QEGWIHHAR HAVACFLITG DLKISKE 385	
MmCPDII 302 ILIWKEISON FCYINPGYDGEESSPEW AKESLNAHEN DYRSHITLE EFERGAGTHOP IM NASQMELL STGKMINETR WILAKHI-E BESEPEKALE 397 ECCEDI 275 LI-WREFYRH LITYHREICKHRFENAT DRY GONSTANLO ANGEGKTGYP IVDAANKQLN STGKMINETR WILAKHI-E DESEPEKALE 360 AnCEDI 284 LA-WREFYRH LITYHRHEICKHRFENAT DRY GONSTANLO ANGEGKTGYP IVDAANKQLN STGKMINERK MILASTL-K LILIDWR 360 TCCEDI 284 LA-WREFYRH LITYHRHDADECPYREIMO QFP GONSTANLO ANGEGKTGYP IVDAANKQLN ETGGMINERK MILASTL-K LILIDWR 360 TCCEDI 284 LL-WREFYRH LITYHRHDAGUENDER FQAFF GONSTANLO ANGEGKTGYP IVDAANKGLN HAGFLSTARA MINASPAYKH KLI GONSTANLO ANGEGKTGYP IVDAANKELS MINASPAYKH KLI	
EccEpi 275 LI-WREFYRH LITYHPELCKHRFFIAMT DRV COSNPAHLQ ANGESKTGYP IVDAANRELN STGRMHRHER MITASFLV-K DLLIDAR 360 AnCPDI 284 LA-WREFYGH ALYHFPELADGPYRSIMQ QFP FORREALF ANTQAQTGYP IVDAANRELT STGRMHRHER MITASFLV-K DLLIDAR 368 TCCEDI 245 LL-WREFYSH LLYHFPEMAREPLDER FQAFP COEDEALFQ ANYEGKTGYP LVDAANRELT ATGELSRAR NNAAQFAVKH LLLIDAR 368 326 LL-WRDFYFR LSIKGGNSIFHLGGRRNV Q	
AncPDI 294 LA-WREFYOH ALIYHFPSLADCEYYRSLMQ OPF "WENREALFT ANTOGATGYP IVDAAMKELT ETGMMIRKER MIVASFLT-K DLLINAR 368 TCCPDI 245 LL-WREFYNH LLYHFPMACREPIDER FOAFF "WENREALFD ANYEKTGYP UVDAAMKELT MAGTLENKRA MINAAPTAVHK 329 AtDASH 326 LI-WRDYFRF LSIKGGKSLFHLGGFRWV QG KNSQDQKLFE SWRDAKTGYP LIDANMKELS TTGFMSNGR QIVCSFLV-R DMGLOAR 410 Dm(6-4) 300 LM-WREFYT VAAAEPNFDR MLGNVYCMQI P QEHPDHLE AWTHGRTGYP FIDAIMROLT QEGWIHLAR HAVACFLTRG DLWISME 385	
TtCPDI 245 LL-WRDFSYH LLYHFPWMAERPLDPR FQAFPWGEDEALFQ AWYEGKTGVP LVDAAMRELH ATGFLSARAR MNAAQFAVKH LLLFAK 329 AtDASH 326 LL-WRDFFRF LSIKGGNSL -FHLGGRRW QG KNSQDQKLFE SWRDAKTGVP LIDANMKELS TTGFMSMRGR QIVCSFLV-R DMGLDWR 410 Dm(6-4) 300 LM-WREFYYT VAAAPPNFDR MLGNVYCMQI P	
AEDASH 326 LI-WRDYFRF LSIKGGNSIFHLAGFRAV QG KNGODOKLFE SWRDAKTGYP LIDANMKELS TTGFNSNGR QIVCSFLV-R DMGLDAR 410 Dm(6-4) 300 LM-WREFYYT VAAAEPNFDR MLGNVYCMQI P DEHPDHLE AWTHGRTGYP FIDAIMROLR QEGWIHHLAR HAVACFLIRG DLWISME 385	
Dm(6-4) 300 LM-WREFYYT VAAAEPNFDR MLGNVYCMQI P	
Dunge-4, 300 FW-AREFILI ANAREANEDK MEGAVICHAL F	
MMACPDII 398 IAICINDRYE LDGRDPNGYA GIAWSIGGVH DRAWGEREVT GKIRYMSY EGCKRKFDV- KLYIEKYS	462
ECCPDI 361 EGERYFMSQL IDGDLAANNG G QWAASTGT DAAPYFRIFNP TTQGEKFDHE GEFIRQWLPE LRDVPGKVVH EPWKWAQKAG VTLDYPQPIV EHKEARVQTL AAYEAARK	469
AnCPDI 369 RGEQFFMQHL VDGDLAANNG GMQWSASSGM DP KPLRIFNP ASQAKKFDAT ATYIKRWLPE LRHVHPKDLI SGEITPIERR GYPAPIVNHN LRQKQFKALY NQLKAAI	475
TtCPDI 330 RCEEAFRHLL LOGDRAVELQ GROWAGGLGV DAAPYFRVFNP VLQGERHDPE GRWLKRWAPE YPSYAPKDPV VDLEEARRRY LRLARD	416
AtDASH 411 MGAEWFETCL LDYDPCSING NWTYGAGVG- NDPREDRYFSI PKQAQNYDPE GEYVAFWLQQ LRRLPKEKRH WPGRLMYMDT VVPLKHGNG	499
Dm(6-4) 386 EGORVEROLL LOODWALWAG NWWULSASAFFH OYFR-VYSPV AFG-KKTDPO GHYLKYVPE LSKYPAGCIY EPWKASLVDO BAYGCVLGTD YPHRIVKHEV VHKENIKEMG AAYKVNEF	

R			(••••••••••••••••••••••••••••••••••••••	α10			@11		0	α12 -	
MmCPDII	244	FIADRLDSYG AL <mark>RN</mark> DPTK-	NMLSNLSP	YLHFGQISSQ	RVVLE	VEKAES	-NPGSKKAFL	DEILIWKEIS	DNFCYYNPGY	DGFESFPS	WAKESLNAHR	337
MaCPDII	242	FLAARLDSYN TL <mark>RN</mark> DPTK-	NALSNLSP	YLHFGQISAQ	R VVLE	VEKAKS	-DPESKKAFL	D <mark>e</mark> ilvwkeia	DNFCYYNPGY	DSFESFPD	WAKKSLNAHR	335
MCPDII	242	FLTNKLDSYS SL <mark>RN</mark> DPTK-	DALSNLSP	YLHFGQISAQ	RVALK	VEKAKA	-DLESKRVFL	D <mark>e</mark> llvrkela	DNFCYYNPFY	DSFDGFPD	WAKKTLNSHR	335
MtCPDII	227	FLREKLECFE RY <mark>RN</mark> DPVK-	NCLSNMSP	YLHFGQISPL	YLALR	AS	-EAGECPEFL	EELIVRRELS	MNFVHYSDSY	SSISCLPE	WAQRTLMDHV	316
AtCPDII	248	FLTKRLKNYS TD <mark>RN</mark> NPIKP	KALSGLSP	YLHFGQVSAQ	RCALE	ARKVRST-	-SPQAVDTFL	EELIVRRELS	DNFCYYQPHY	DSLKGAWE	WARKSLMDHA	343
<i>Os</i> CPDII	260	FLTKRIKSYE TD <mark>RN</mark> DPTKP	RALSGLSP	YLHFGHISAQ	RCALE	AKKCRHL-	-SPKSVDAFL	EELVVRRELA	DNFCYYQPQY	DSLSGAWE	WARKTLMDHA	355
DsCPDII	300	LTTPRIAQYH VK <mark>RN</mark> DPSCT	TGLSNLSP	YLHFGQLSAQ	RAALE	ASKLRSR-	-HREAVDRYL	EELIVRRELA	DNFCEHCPDY	DKLVPGTAYD	WALKSLEKHK	397
DmCPDII	329	FCSRRLRHFN DKRNDPTA-	DALSGLSP	WLHFGHISAQ	RCALE	VQRFRGQ-	-HKASADAFC	EBAIVRRELA	DNFCFYNEHY	DSLKGLSS	WAYQTLDAHR	423
XLCPDII	332	FISERLKHFN SD <mark>RN</mark> NPNQ-	NALSNLSP	WFHFGQLSVQ	RAILE	VQKYRSK-	-FKESVDSFV	EEAVVRRELA	DNFCFYNKNY	DKIEGAYD	WAKNTLKDHA	426
PtCPDII	310	FIAERLPYFG SD <mark>RN</mark> NPNK-	DALSNLSP	WFHFGQVSVQ	RAILE	VQKHRSR-	-YPDSVTNFV	EEAVVRRELA	DNFCFYNKNY	DKLEGAYD	WAQTTLRLHA	404
CaCPDII	321	FIDQRLRLFA TH <mark>RN</mark> NPNY-	DALSHLSP	WIHTGQLSAQ	RVVKQ	VKREKN	-ASESVASFI	EELVVRRELA	DNFCFYNPSY	DNISGAYD	WAKKTLQDHA	414
FVCPDII	245	FIKNRLPSYD AD <mark>HN</mark> NPTC-	DALSNLSP	WLHFGHVSAQ	RVALE	VLKCIRE-	-SKKNVETFI	DEIIVRRELS	DNFCYYNKHY	DSIQSTHS	WARKTLEDHI	339
HOCPDII	237	FLHNKIKDYH EY <mark>RN</mark> DPVK-	NWISNMSP	YLHFGQVSPL	HLIIK	GNNYCKKHE-	-IDKGFKEFF	EELVIRRELS	FNFVYYNPDY	DSIKSLPD	WAKKTLKEHE	333
DdCPDII	237	FIRVRMHRYA AE <mark>RN</mark> NPLM-	PVLSHLSP	YLHFGMLSAQ	RAVLEVMQAQ	GTAAGGT-	-YGEGAAAFV	EELVVRRELA	DNFCWYEPSY	DSVEAFPD	WALKTLDRHR	336
CfCPDII	227	FLEKRLSSYA EL <mark>RN</mark> DPNS-	GVLSNLSP	YLHFGQISAQ	YIALR	VSESRM	-PDESRSAFL	EELIVRRELS	DNYCFYNDRY	DSFDGSPT	WAKESLMNHR	320
GSCPDII	230	FLEDGLAGYA TR <mark>RN</mark> NPAV-	MGQSGLSP	WLHFGQLSAQ	RVAQ	AAFAAAA-	-PIESRDAFL	EELIVRRELA	DNFCYYNDAY	DRFDGFPE	WAQRTLNRHR	323
			-(«14		@15		@16		(1)			8
							TATOL	1000000				
MmCPDII	338	NDVRSHITTL EEFEAGRTH	D PLWNASQMEL	LOTGENHUGYN	RMIWARKILE	WSESPERALE	UNICLNDRIE	LDGRDPNGIA	GIAWSIGGVH	DRAWGEREVT	GKIRIMSIEG	447
MACPDII	336	DODGULERI PELERCORY	D PLWNASQIEL	LRIGKMHSIM	RMIWARKILE	WSESPERALE	VAICLNDRIE	LDGRDPNGIA	GIAWSIGGLH	DRAWREREVI	GKIRIMSIEG	445
MDCPDII	336	RDQRSHIFIL EELEIGKII	D PLANASQIEL	NTECKNUCYM	RMINARKILE	WEDUDADAYD	TATTENDATE	LDGRDPNGIA	GIAWSIGGVH	DRAWQEREIF	GKIRIMSIEG	445
AtCPDII	311	SURPRETEISL RELESASIN	D PIWNAAQQEM	VYOCKNHOPM	PMYMOKKTLE	WTECPEENIS	TSTYLNNKYE	TOGROPAGEA	GOMWSTCGVH	DOGWKERDVE	GKTRYMNUNG	425
OCCEDIT	356	ADEREHINTE FOLENAETH	D DIWNASQUEM	VHHCKMHCFM	DWAMPKKITE	WTSCPEEALS	TAIVINDEVE	TOGROPSCVV	GCMWSICGIH	DOGWKERPVE	GKIRYMNYAG	455
DeCPDII	398	RDPRPITYTE OOLESGHTG	D DIWNAGOMEL	VROGKMHGYI	RMYWAKKILE	WSGTPEEAVE	NATYINDKWS	LDGRDPSGYT	GVMWSVAGVH	DRAWIDRPIY	GKIRVMTYDG	507
DmCPDTT	424	KDKRDPCYSL EELEKSLTY	D DLWNSAOLOL	VREGKMHGEL	RMYWAKKILE	WTATPEHALE	YATLINDKYS	LDGRDPNGYV	GCMWSIGGVH	DMGWKERAIF	GKVRYMNYOG	533
XICPDII	427	KDKRTHLYTL EKLEAGKTH	D PLWNAAOLOM	VHEGKMHGFL	RMYWAKKILE	WTSSPEEALH	FSLYLNDRYE	LDGRDPNGYV	GCMWSICGIH	DOGWAERAVE	GKIRYMNYOG	536
PtCPDII	405	KDKRPHLYSL EOLESGKTH	D PLWNAAOMOT	VKEGKMH <mark>G</mark> FL	RMYWAKKILE	WTRSPEEALE	FAIYLNDRFO	LDGWDPNGYV	GCMWSICGIH	DOGWAEREIF	GKIRYMNYAG	514
CaCPDII	415	KDSROYLYTK EQLENAKTH	D OLWNAAOROL	VSEGKMH <mark>G</mark> FL	RMYWAKKILE	WTASPEEALS	IAIYLNDRLS	LDGCDPN <mark>G</mark> YV	GCMWSICGIH	DQGWAERPIF	GKIRFMNYAG	524
FvCPDII	340	NDPRKYIYSI KQLEKAETH	D PLWNASQMQM	VREGKMHSFL	RMYWAKKILE	WTRTPEDALS	YSIYLNNKYE	LDGTDPN <mark>G</mark> YV	GCMWSICGLH	DRAWKERPIF	GKIRYMNYES	449
HOCPDII	334	NDTREFSTSL QELEDAKTH	D PY <mark>W</mark> NAAQKEL	LLTGKIH <mark>G</mark> YM	RMINGKKILE	WTSSPDLAYK	YALYL <mark>N</mark> NKYA	ldgrdpn <mark>g</mark> fa	GVAWCFG-KH	DRPWPGCNIF	GKVRYMSSGG	442
DdCPDII	337	ADRRPYL <mark>Y</mark> DE QQLEKARTH	D PL <mark>W</mark> NAAQQEM	VLTGKMH <mark>G</mark> YM	RMYWAKKILE	WTESPEEALR	IVIRONDRWS	ldgrdsn <mark>g</mark> ya	GAAWSVGGVH	DRPWREREVF	GTIRFMSYNG	446
						WCECSPORER	TRMAT	TDCDDDNCVB	CUBRCTCCUU	DDDWFFDDWV	CRIDYMNBCC	
CfCPDII	321	NDHREYLYTA DEFATAKTH	D RL <mark>W</mark> NAAQLEL	VTTGKIH <mark>G</mark> YM	RMYWAKKILE	"STRATCHLE	IAPALADRIA	LDGKDFNGIA	GVANDIGGVH	DIVENTIAL	GUTUINNAGG	430
<i>Cf</i> CPDII <i>Gs</i> CPDII	321 324	NDHREYL Y TA DEFATAKTH HDPRPQC <mark>Y</mark> EH DVLEQGQTH	D RL <mark>W</mark> NAAQLEL D SL <mark>W</mark> NAAQLEM	VTTGKIH G YM VRWGRMH <mark>G</mark> YL	RMYWAKKILE	WTSSPEDALM	IAIQLNDRIQ	LDGRDPN G YA	GIAWSIGGVH	DRPWAERPVF	GTIRFMSRDG	430

		••••	$\cdots \cdots \cdots $	$\cdots \mid \cdots \mid$	$\cdots \cdots \mid \cdots \mid$	$\cdots \cdots \mid \cdots \mid$
MmCPDII	448	CKRKFDVKLY	IEKYSAL			
MaCPDII	446	CKRKFNVKLY	IAKYSAL			
MbCPDII	446	SKRKFDVKSY	IDKYSAL			
MtCPDII	426	LKRKFRIDEY	VDRIRGLMDE			
AtCPDII	454	CKRKFNVDSY	ISYVKSLVSV	TKKKRKAEEQ	LTRDSVDPKI	$\mathbb{T} \mathbb{I} \mathbb{V} = = = = = = = = = = = = = = = = = = =$
<i>Os</i> CPDII	466	CKRKFDVDAY	ISYVKRLAGQ	SKK-RNAEES	PNPVVKLSKS	QH
DSCPDII	508	CKGKFDVPAY	VAYVESLGRA	RQ		
DmCPDII	534	CRRKFDVNAF	VMRYGGKVHK	KK		
XICPDII	537	CKRKFDVAQF	ERRYHPKKFS	Q		
PtCPDII	515	CKRKFDVAEF	ERKISPAD			
CaCPDII	525	CKRKFDVAQF	ERKYTAVKEN	SNKDSKKSSS	KN	
FvCPDII	450	SKKKFDVAVF	IQKYN			
<i>Ho</i> CPDII	443	LKRKFKIDLY	LKRIHNLEEA	SHVG		
DUCPDII	447	ARSKFDVDGY	VAAVAALENII	PVPAPVSRCG	GRRKPAQGLL	L
CfCPDII	431	CARKFDVARY	IARFEEKKSM			
GSCPDII	434	CRRKFDTDAY	ERRVIISP	ATCAGIALCK		

Figure S4. Multiple sequence alignments of the catalytic subdomain of photolyases. (A) Structure-based sequence alignment of *Mm*CPDII with the class I photolyases from *E. coli* (1DNP), *A. nidulans* (1TEZ), *T. thermophilus* (1IQR), the DASH-like cryptochrome 3 from *A.*

thaliana (2VTB) and the 6-4 photolyase from D. melanogaster (3CVV). (B) Multiple sequence alignment for class II photolyases using ClustalW2 (Larkin et al, 2007). Numbering refers to the corresponding UniProtKB entries. Secondary structure motifs for the *M. mazei* photolyase were assigned by STRIDE (Heinig & Frishman, 2004). Conserved elements are highlighted in green for class II photolyases and in orange for other subclasses: (1) Tryptophans (red) of the dyad (MmCPDII) and triad (other photolyases), (2) surface-exposed residues involved in electron transfer pathway (purple), (3) stabilizing asparagine of the neutral radical state of the catalytic cofactor (white) and class II conserved glycine (brown), (4) glutamate at 5'-thymidine and asparagine at 3'-thymidine (blue), (5) alternative asparagine at 3'-thymidine (MmCPDII N257, yellow) and stabilizing arginine (MmCPDII R256, yellow). The missing C-terminal extension for class II photolyases is highlighted in pale blue. Abbreviations used for class II photolyases: Methanosarcina mazei (MmCPDII), Methanosarcina acetivorans (MaCPDII), Methanosarcina barkeri (MbCPDII), Methanobacterium thermoautotrophicum (MtCPDII), Arabidopsis thaliana (AtCPDII), Oryza sativa (OsCPDII), Dunaliella salina (DsCPDII), Drosophila melanogaster (DmCPDII), Xenopus laevis (XlCPDII), Potorous tridactylus (PtCPDII), Carassius auratus (CaCPDII), Fowlpox virus (FvCPDII), Halothermothrix orenii (HoCPDII), Desulfovibrio desulfuricans (DdCPDII), Chlorobium ferrooxidans (CfCPDII), Geobacter sulfurreducens (GsCPDII).

References

Heinig M, Frishman D (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res* **32**(Web Server issue): W500-502

Hill AV (1910) The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *Journal of Physiology* **40**(Suppl): i-vii

Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJA, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C (2009) InterPro: the integrative protein signature database. *Nucleic Acids Res* **37**: D211-D215

Husain I, Griffith J, Sancar A (1988) Thymine dimers bend DNA. *Proc Natl Acad Sci U S A* **85**(8): 2558-2562

Jorns MS, Sancar GB, Sancar A (1985) Identification of oligothymidylates as new simple substrates for Escherichia coli DNA photolyase and their use in a rapid spectrophotometric enzyme assay. *Biochemistry* **24**(8): 1856-1861

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**(21): 2947-2948

Lassmann T, Sonnhammer ELL (2005) Kalign - an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics* 6

Mees A, Klar T, Gnau P, Hennecke U, Eker AP, Carell T, Essen LO (2004) Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair. *Science* **306**(5702): 1789-1793

Park H, Zhang K, Ren Y, Nadji S, Sinha N, Taylor JS, Kang C (2002) Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc Natl Acad Sci U S A* **99**(25): 15965-15970

Pearlman DA, Holbrook SR, Pirkle DH, Kim SH (1985) Molecular models for DNA damaged by photoreaction. *Science* **227**(4692): 1304-1308

Pokorny R, Klar T, Hennecke U, Carell T, Batschauer A, Essen LO (2008) Recognition and repair of UV lesions in loop structures of duplex DNA by DASH-type cryptochrome. *Proc Natl Acad Sci U S A* **105**(52): 21023-21027

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**(4): 406-425

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**(8): 1596-1599

UniProtConsortium (2009) The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res* **38**(Database issue): D142-148

Supplementary Tables

Suppl. Table I. Preliminary crystallographic statistics of *M. barkeri* class II photolyase.

data collection & processing	MbCPDII
uata concetion & processing	ID14-2
X-ray source	ESRE Grenoble France
detector	ADSC 04
wavelength (Å)	0.9330
space group	$P3_1$
cell dimensions $(a h c Å)$	119 41, 119 41, 100 26
resolution (Å)	51 71-2 30 (2 42-2 30)
total reflections	205222
multiplicity	2.9 (2.9)
unique reflections	71031
$R_{\text{marge}}(\%)$	7.0 (48.9)
completeness (%)	99.9 (100.0)
$I/\sigma(I)$	11.0 (2.3)
mosaicity (°)	0.23
Wilson B-factor ($Å^2$)	41.5
× /	
refinement statistics	
resolution (Å)	51.30-2.30
$R_{\text{factor}}, R_{\text{free}}$ (%)	18.7, 23.2
reflections (working, test set)	70082, 897
completeness for range (%)	99.9
r.m.s.d. from ideal:	
bond lenghts (Å)	0.012
bond angles (°)	1.376
total number of atoms	10508
mean <i>B</i> value ($Å^2$)	19.6

Mutant	Primer	Expression condition
<i>Mm</i> CPDII W388F	5'- GGGCAAAAAAAATTCTGGAATTCAGCGAATCTCCCG-3' 5'- CGGGAGATTCGCTGAATTCCAGAATTTTTTTGCCC-3'	20 °C (24h)
<i>Mm</i> CPDII W360F	5'- CACATGACCCACTCTTCAACGCTAGCCAGATGGAACTTC-3' 5'-GAAGTTCCATCTGGCTAGCGTTGAAGAGTGGGTCATGTG-3'	25 °C (24h)
<i>Mm</i> CPDII W381F	5'-CACGCGCATGTACTTCGCAAAAAAAATTCTAGAATGGAGCG-3' 5'- CGCTCCATTCTAGAATTTTTTTTGCGAAGTACATGCGCGTG-3'	15 °C (48h)
<i>Mm</i> CPDII Y345F	5'-GTGAGGAGTCATATCTTCACTCTAGAAGAGTTCGAAGC-3' 5'-GCTTCGAACTCTTCTAGAGTGAAGATATGACTCCTCAC-3'	20 °C (24h)
<i>Mm</i> CPDII Y380F	5'-GCACGGTTACACGCGTATGTTCTGGGCAAAAAAAATTCTGG-3' 5'-CCAGAATTTTTTTGCCCAGAACATACGCGTGTAACCGTGC-3'	20 °C (24h)
<i>Mm</i> CPDII N403D	5'-GAAATTGCAATCTGCCTGGACGATCGGTATGAACTTGACG-3' 5'-CGTCAAGTTCATACCGATCGTCCAGGCAGATTGCAATTTC-3'	15 °C (24h)
<i>Mm</i> CPDII N403L	5'-GAAATTGCAATCTGCCTGCTCGATCGGTATGAACTTGACG-3' 5'-CGTCAAGTTCATACCGATCGAGCAGGCAGATTGCAATTTC-3'	15 °C (24h)
<i>Mm</i> CPDII N403A	5'-GAAATTGCAATCTGCCTGGCCGATCGGTATGAACTTGACG-3' 5'-CGTCAAGTTCATACCGATCGGCCAGGCAGATTGCAATTTC-3'	15 °C (24h)

Suppl. Table II. Details of mutagenesis of *M. mazei* class II photolyase.