**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 *Mm*CPDII 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 Ni-

NTA 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 *Mm*CPDII**

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 *Mm*CPDII•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 CPD-lesion, 5'-IRDye700-AAAATGCTGGATGTCGAGGTGTAAT<>TAATGTGGAGCTGTAGGTCGTAAAA-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'-TTTTACGACCTACAGCTCCACATTAATTACACCTCGACATCCAGCATTTT-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**

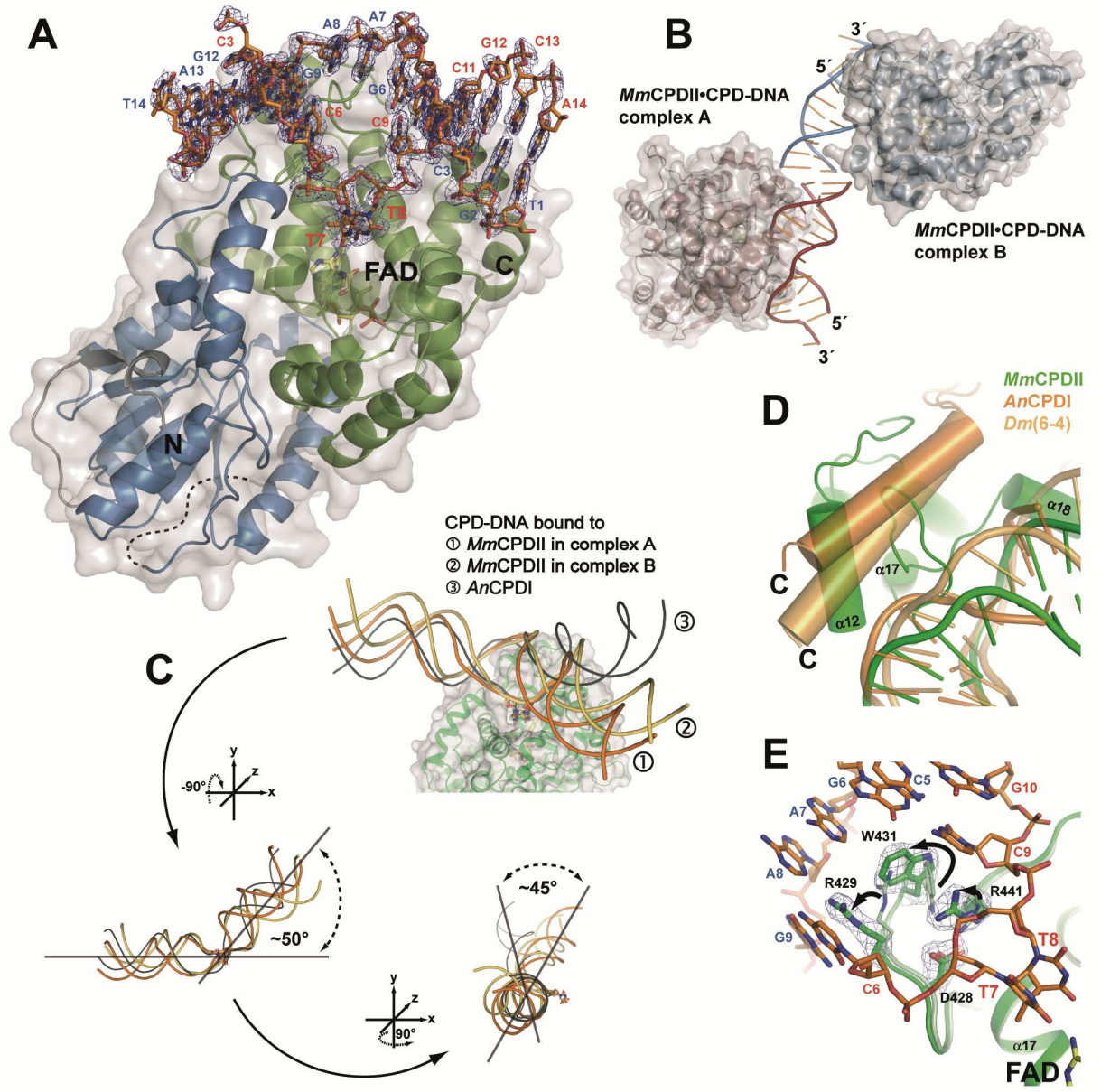


**Figure S1. *M. mazei* class II photolyase complexed to CPD-DNA.** Electron density (SIGMAA-weighted $2F_{obs}$-$F_{calc}$) is contoured at 1 σ. (A) The 14mer duplex DNA comprising the chemically synthesized CPD-lesion is bound to the catalytic domain of the photolyase. The thymine dimer (T7<>T8) is 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 α17 and α18. (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.

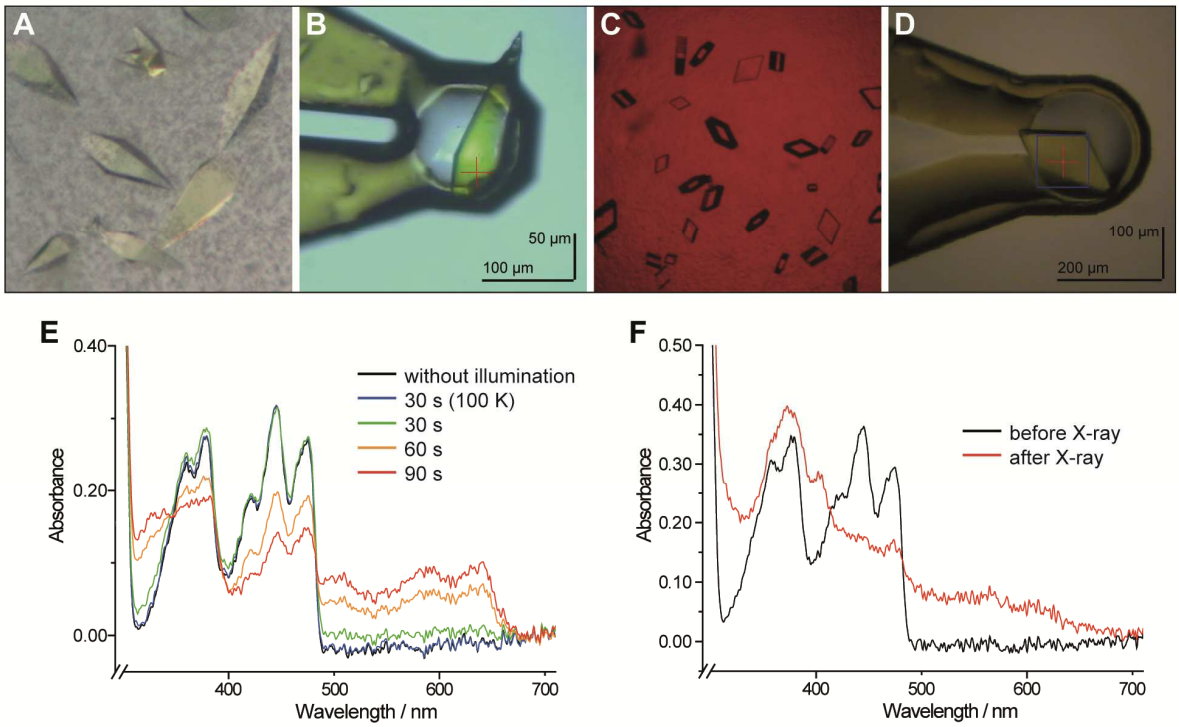


**Figure S2. *Mm*CPDII crystals and *in crystallo* UV/Vis spectroscopy.** (A and B) Crystals of *M. mazei* class II photolyase and (C and D) in complex with CPD-DNA. Images of crystals

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 illumination at 450 nm (green, orange and red) was only achieved, when the crystal was illuminated at room temperature. (F) X-ray radiation (flux: ~5.0•$10^{10}$ 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.

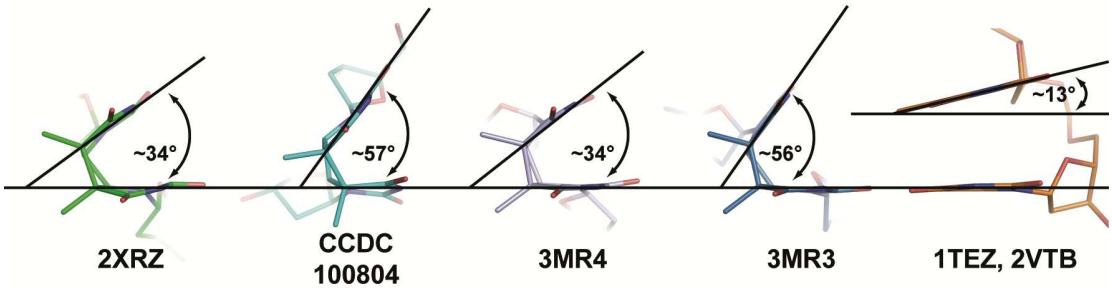


**Figure S3. Angular analysis of CPD-lesions.** Given angles between the 5´- and 3´-thymidine base planes were calculated using two normal straight lines (2XRZ: *Mm*CPDII•CPD-DNA; CCDC100804: Synthetic CPD-lesion; 3MR3, 3MR4: Human DNA polymerase η•CPD-DNA; 1TEZ: *An*CPDI•CPD-DNA; 2VTB: *At*DASH•CPD-DNA).


**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 (*Mm*CPDII) 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 (*Mm*CPDII N257, yellow) and stabilizing arginine (*Mm*CPDII R256, yellow). The missing C-terminal extension for class II photolyases is highlighted in pale blue. Abbreviations used for class II photolyases: *Methanosarcina mazei* (*Mm*CPDII), *Methanosarcina acetivorans* (*Ma*CPDII), *Methanosarcina barkeri* (*Mb*CPDII), *Methanobacterium thermoautotrophicum* (*Mt*CPDII), *Arabidopsis thaliana* (*At*CPDII), *Oryza sativa* (*Os*CPDII), *Dunaliella salina* (*Ds*CPDII), *Drosophila melanogaster* (*Dm*CPDII), *Xenopus laevis* (*Xl*CPDII), *Potorous tridactylus* (*Pt*CPDII), *Carassius auratus* (*Ca*CPDII), Fowlpox virus (FvCPDII), *Halothermothrix orenii* (*Ho*CPDII), *Desulfovibrio desulfuricans* (*Dd*CPDII), *Chlorobium ferrooxidans* (*Cf*CPDII), *Geobacter sulfurreducens* (*Gs*CPDII).

## 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** | ***Mb*CPDII** |
|---|---|
| X-ray source | ID14-2<br>ESRF, Grenoble, France |
| detector | ADSC Q4 |
| wavelength (Å) | 0.9330 |
| space group | $P3_1$ |
| cell dimensions (*a*,*b*,*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 |
| $R_{merge}$ (%) | 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 |
| $R_{factor}$, $R_{free}$ (%) | 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 *B* value (Å$^2$) | 19.6 |

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

| Mutant | Primer | Expression condition |
|---|---|---|
| *Mm*CPDII W388F | 5'- GGGCAAAAAAAATTCTGGAATTCAGCGAATCTCCCG-3'<br>5'- CGGGAGATTCGCTGAATTCCAGAATTTTTTTTGCCC-3' | 20 °C (24h) |
| *Mm*CPDII W360F | 5'- CACATGACCCACTCTTCAACGCTAGCCAGATGGAACTTC-3'<br>5'-GAAGTTCCATCTGGCTAGCGTTGAAGAGTGGGTCATGTG-3' | 25 °C (24h) |
| *Mm*CPDII W381F | 5'-CA**C**GCGCATGTACTTCGCAAAAAAAATTCTAGAATGGAGCG-3'<br>5'- CGCTCCATTCTAGAATTTTTTTTGCGAAGTACATGCGCGTG-3' | 15 °C (48h) |
| *Mm*CPDII Y345F | 5'-GTGAGGAGTCATATCTTCACTCTAGAAGAGTTCGAAGC-3'<br>5'-GCTTCGAACTCTTCTAGAGTGAAGATATGACTCCTCAC-3' | 20 °C (24h) |
| *Mm*CPDII Y380F | 5'-GCACGGTTACACGCGTATGTTCTGGGCAAAAAAAATTCTGG-3'<br>5'-CCAGAATTTTTTTTGCCCAGAACATACGCGTGTAACCGTGC-3' | 20 °C (24h) |
| *Mm*CPDII N403D | 5'-GAAATTGCAATCTGCCTGGACGATCGGTATGAACTTGACG-3'<br>5'-CGTCAAGTTCATACCGATCGTCCAGGCAGATTGCAATTTC-3' | 15 °C (24h) |
| *Mm*CPDII N403L | 5'-GAAATTGCAATCTGCCTGCTCGATCGGTATGAACTTGACG-3'<br>5'-CGTCAAGTTCATACCGATCGAGCAGGCAGATTGCAATTTC-3' | 15 °C (24h) |
| *Mm*CPDII N403A | 5'-GAAATTGCAATCTGCCTGGCCGATCGGTATGAACTTGACG-3'<br>5'-CGTCAAGTTCATACCGATCGGCCAGGCAGATTGCAATTTC-3' | 15 °C (24h) |