bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Dimerisation of European robin cryptochrome 4a

Maja Hanić[1#], Lewis M. Antill[2,3#], Angela S. Gehrckens[4#], Jessica Schmidt[5], Katharina Görtemaker[6], Rabea Bartölke[5], Tarick J. El-Baba[4,7], Jingjing Xu[5], Karl W. Koch[6,8], Henrik Mouritsen[5,8], Justin L. P. Benesch[4,7], P. J. Hore[4], and Ilia A. Solov'yov[1,8,9*]

[1]*Institute of Physics, Carl von Ossietzky University of Oldenburg, Carl-von-Ossietzky Straße 9-11, 26129, Oldenburg, Germany*

[2]*Graduate School of Science and Engineering, Saitama University, 255 Shimo-okubo, Sakura Ward, Saitama 338-8570, Japan*

[3]*JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

[4]*Department of Chemistry, University of Oxford, Physical & Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, UK*

[5]*Department of Biology and Environmental Sciences, Carl von Ossietzky University of Oldenburg, Carl-von-Ossietzky Straße 9-11, 26129, Oldenburg, Germany*

[6]*Department of Neuroscience, Division of Biochemistry, Carl von Ossietzky University of Oldenburg, D-26111, Oldenburg, Germany*

[7]*Kavli Institute for NanoScience Discovery, Dorothy Crowfoot Hodgkin Building, University of Oxford, OX1 3QU*

[8]*Research Center for Neurosensory Sciences, Carl von Ossietzky University of Oldenburg, Carl-von-Ossietzky Straße 9-11, 26111, Oldenburg, Germany*

[9]*Center for Nanoscale Dynamics (CENAD), Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstr. 114-118, 26129 Oldenburg, Germany*

*Author for correspondence: ilia.solovyov@uni-oldenburg.de

[#]These authors contributed equally

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Abstract

Homo-dimer formation is important for the function of many proteins. Although dimeric forms of cryptochromes (Cry) have been found by crystallography and were recently observed *in vitro* for European robin Cry4a, little is known about the dimerisation of avian cryptochromes and the role it could play in the mechanism of magnetic sensing in migratory birds. Here we present a combined experimental and computational investigation of the dimerisation of robin Cry4a resulting from covalent and non-covalent interactions. Experimental studies using native mass spectrometry, mass spectrometric analysis of disulphide bonds, chemical cross-linking and photometric measurements show that disulphide-linked dimers are routinely formed, the most likely cysteines being C317 and C412. Computational modelling and molecular dynamics simulations were used to generate and assess a number of possible dimer structures. The relevance of these findings to the proposed role of Cry4a in avian magnetoreception is discussed.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Introduction

Oligomerisation of proteins can change their structural stability, activity, and mechanisms of action[1,2] and is important in numerous processes in living organisms. Examples from the cryptochrome (Cry) protein family include a tetrameric structure of *Arabidopsis thaliana* Cry2 (*At*Cry2) important in photoactivation[3], a proposed dimer of L-Cry involved in the reproduction cycle of the bristle worm, *Platynereis dumerilii*[4], and dimerisation of *At*Cry1[5,6].

In the investigation reported here, we focus on Cry, a protein found in all biological kingdoms[7–12]. Crys are structurally similar to light-activated DNA photolyases[7,12,13] and consist of two main structural regions: a highly conserved N-terminal photolyase homology region (PHR) and the intrinsically disordered C-terminal tail (CTT)[14]. Most Crys have a flavin adenine dinucleotide cofactor (FAD) that absorbs UV-blue light, non-covalently bound in the PHR domain[7]. Crys can exist in the so-called dark-state, containing the fully oxidised form of the FAD, or in higher energy states formed by photo-excitation of the FAD followed by the formation of radical pairs, involving a triad or tetrad of tryptophan residues, which have been proposed as mediators of the mechanism of magnetoreception in migratory songbirds[15–20]. Three different Cry-genes, Cry1[21–26], Cry2[21,27], and Cry4[28–32] exist in most bird species, each with at least one splice-variant[22,33,34]. Several of these cryptochromes are found in the birds' retina, where they can easily be light-activated[21,24,25,27,29,32,34]. Light-activated Cry is believed to form a signalling state, via a conformational change in the CTT, which has been the subject of computational studies of European robin Cry4a (*Erithacus rubecula*, *Er*Cry4a)[35] and pigeon Cry4a (*Columba livia*, *Cl*Cry4a)[36] and experimental investigations of Crys from the fruit fly (*Drosophila melanogaster*, *Dm*Cry)[37], chicken (*Gallus gallus*, *Gg*Cry4a)[38] and *Arabidopsis* (*At*Cry1)[39]. The exact process by which Cry combines photo-excitation with detection of the Earth's magnetic field to form a long-lived signalling state is unknown but could conceivably involve dimerisation. The importance of homo-oligomerisation of plant Crys *in vivo* is clear[40,41] and optogenetic studies of *At*Cry2 show that the full-length protein undergoes light-induced oligomerisation and that functionality is lost in a truncated form that lacks the CTT domain[42].

For a signal transduction cascade to function, Cry must interact with other proteins. Wu *et al.* recently listed six proteins as candidate *Er*Cry4a interaction partners[43]. A subsequent detailed biochemical investigation of *Er*Cry4a and a cone-specific G-protein from European

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

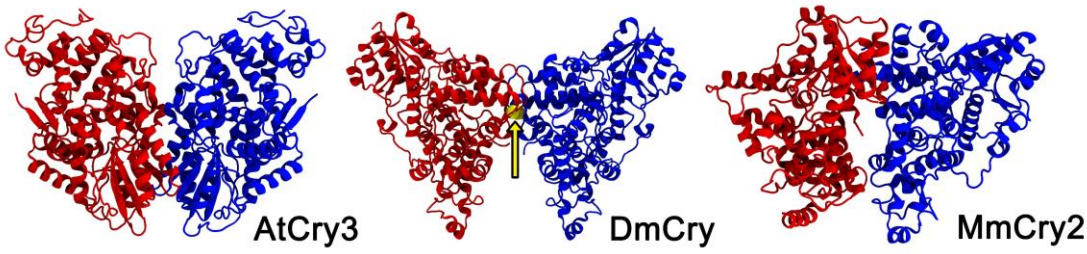

Figure 1. Structures of cryptochrome dimers from the Protein Data Bank (PDB): non-covalent *At*Cry3 dimer (PDB: 2J4D[3]), covalent *Dm*Cry dimer (PDB: 4GU5[28]) (the arrow indicates the C296-C296 disulphide bond) and non-covalent *Mm*Cry2 dimer (PDB: 6KX8[29]).

robin demonstrated that these two proteins interact directly with each other[44]. Dimers of *Er*Cry4a have also been reported[15], but is unclear whether/how a monomer-dimer equilibrium might be involved in the downstream signalling process. Within the Cry family there are reports of crystallographic dimeric asymmetric units, including *Dm*Cry[45] and mouse (*Mus musculus*) *Mm*Cry1 and *Mm*Cry2[46] (Fig. 1). A study of full-length *Dm*Cry noted an intermonomer disulphide bond, but deemed it to be an artefact of the crystallisation method[47]. To date, there has been little work on the dimers of animal Crys, and their function, if any, is unknown.

To learn more about *Er*Cry4a dimerisation and its possible role in magnetoreception[15,18], we have explored a variety of candidate structures, including covalently and non-covalently linked forms of the full-length and truncated protein, using a combination of experimental and computational methods to identify potential *Er*Cry4a dimers. Native mass spectrometry (MS), mass photometry (MP), and gel electrophoresis of wild-type (WT) and mutant proteins were used to establish the presence and nature of the dimers, while chemical cross-linking followed by mass spectrometry (XL-MS) provided information about the relative orientation of the monomer units. A combination of molecular docking and molecular dynamics (MD) techniques provided model structures for comparison with the experimental data.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Methods

Experimental Methods

Protein expression and purification

WT *Er*Cry4a (GenBank: KX890129.1) was cloned, expressed and purified as described by Xu *et al.*[15] with the following modifications. The LB media contained 10 g L^{-1} yeast extract and the expression time was 44 h instead of 22 h. The mutants were generated using the Q5 site-directed mutagenesis kit (New England Biolabs, Ipswich, MA, USA), and plasmids were confirmed by Sanger sequencing (LGC Genomics). For the C412A mutant and the 5-cysteine mutant (C68A+C73A+C116A+C189A+C317A), an *E. coli* K12 codon-optimised version of *Er*Cry4a was used, generated by Eurofins Genomics (Ebersberg, Germany). Tables 1 and 2 summarise the primers used for each mutation. Mutant proteins were expressed and purified in the same way as the WT protein with the following exception: double, triple and 5-cysteine mutants were expressed in BL21-CodonPlus(DE3)-RIPL *E. coli* cells (Agilent, Santa Clara, CA, USA) in the dark, starting with a 30 mL culture grown overnight at 30 °C and 250 rpm, a preculture inoculated to an OD_{600} of 0.05 for 6–6.5 h at 37 °C, and a main culture inoculated with the preculture to an OD_{600} of 0.5 for about 3 h at 37 °C until the OD_{600} reached 0.6, when the shakers were set to 15 °C and 160 rpm. After about 45 min, at an OD_{600} of 0.9 to 1.0, protein expression was induced with 5 µM isopropyl β-D-1-thiogalactopyranoside. Cell harvest, lysis and purification using Ni-NTA agarose columns and anion exchange chromatography were all performed as described in Xu *et al.*[15] Purified protein samples were either used directly for photometric cysteine exposure measurements or buffer-exchanged, to remove reducing agents, into 20 mM Tris, 250 mM NaCl, and 20% glycerol using Sephadex G 25 in PD10 desalting columns (Cytiva, Uppsala, Sweden). Protein samples for XL-MS were produced in insect cells essentially as described previously[44] with the additional step of removing the His-tag. To cleave the His-tag from the Ni NTA eluted protein, samples were buffer-exchanged into 20 mM Tris, 100 mM NaCl, 2 mM *DL*-dithiothreitol, DTT, using Sephadex G 25 in PD10 desalting columns (Cytiva) and the His-tag was cut off overnight at 4 °C by the addition of 10 U AcTEV Protease (Thermo Fisher). *Er*Cry4a was further purified using anion-exchange chromatography as previously described[44]. All purified protein samples were concentrated to ~3.5 mg mL^{-1}.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Samples were snap-frozen in liquid nitrogen and stored at –80 °C until shipment to Oxford on dry ice.

Table 1. Mutations introduced in the non-codon-optimised *Er*Cry4a.

Mutation	Forward primer	Reverse primer
C116A	CAACATCCAGGCCCTGGGGGCAG	GCCTCCATCTCCTTGTAAAATG
C313A	GAACCCCATCGCCCTCCAGATCTGC	CCAGCCATCTGGGTGAAG
C317A	CCTCCAGATCGCCTGGTACAAGGATGCAG	CAGATGGGGTTCCCAGCC
C313A+C317A	GATCGCCTGGTACAAGGATGCAGAGAG	TGGAGGGCGATGGGGTTCCCAGCCAT

Table 2. Mutations introduced in the codon-optimised *Er*Cry4a.

Mutation	Forward primer	Reverse primer
C68A+C73A	GGATCGGCCCTGCTTGTGATTCAGGGTG	AAGCTGGGCCAGGTTCTTATGCAGATCC
C116A	GAACATTCAGGCCTTGGGTGCAG	GCCTCCATCTCCTTGTAG
C189A	TCTTGCCGAAGCCTATCGTGTTCCGTTAC	TCCGGATCTGGAGCCGAA
C317A	TCTGCAGATTGCCTGGTATAAAGACGCAGAAC	CAGATCGGATTGCCCGCC
C412A	CCGGATTTTCGCCCCTGTACGCTTTG	GTGTACTGGTGAAAGAAGG

Native Mass Spectrometry and SDS-PAGE

Native MS and sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) were used to investigate the dimeric nature of *Er*Cry4a. The *Er*Cry4a samples used for these experiments were shipped overnight to Oxford on dry ice from the production laboratory at the University of Oldenburg, in a buffer containing 20 mM Tris, 250 mM NaCl, and 20% glycerol at pH 8. The samples were kept at –80 °C until they were thawed on ice and buffer-exchanged into 200 mM ammonium acetate, pH 8, for the native MS measurements using Zeba™ Micro Spin desalting columns with a molecular weight exclusion limit of 40 kDa (ThermoFisher Scientific). All samples included an N-terminal 10×His-tag used for purification of the proteins.

Measurements were performed in Oxford using in-house gold-plated capillaries on a Q Exactive™ mass spectrometer in positive ion mode. The samples were sprayed using nanoflow electrospray with a source temperature of 150 °C and a capillary voltage of 1.0 kV[48]. The higher-energy C-trap dissociation (HCD) cell voltage was 5 V, in-source trapping was set to –200 V to help with the dissociation of small ion adducts, and the noise threshold

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

parameter was 3. Ion transfer optics and voltage gradients throughout the instruments were adjusted for optimum protein transmission.

For SDS-PAGE gels, the samples were covered in the dark for 20 min at room temperature at a concentration of ~8 µM. 3.2 µg of each sample were then mixed with diluted NuPAGE LDS sample buffer (4X) (Invitrogen, Waltham, MA, USA) before being loaded on NuPAGE Novex 4-12% bis-Tris gels, 1.0 mm × 10 well (Invitrogen) and run with NuPAGE MES SDS running buffer (20X) (Invitrogen). Samples of proteins without a His-tag, containing 2 mM DTT (Sigma-Aldrich) were shipped to Oxford and kept in the light while being incubated with different concentrations of DTT for 20 min at room temperature prior to being submitted to the SDS-PAGE gel. Protein bands were visualised by Coomassie blue staining using QuickBlue Protein Stain (LuBio, Zürich, Switzerland) and accompanied by either a SeeBlue^{TM} Plus2 Pre-stained Protein Ladder (Invitrogen) or a PageRuler Prestained Protein Ladder (Thermo Scientific, Waltham, MA, USA). The software GelAnalyzer 19.1[49] was used for evaluation of the gel bands.

Mass photometry

All samples were shipped as described above except for the *Er*Cry4a WT sample used in Fig. 3 which was shipped with the addition of 2 mM DTT to limit higher order oligomerisation during shipping. This sample was diluted to a concentration of ~44 nM in 200 mM ammonium acetate, pH 8.3, prior to measurement. All other samples were buffer-exchanged into 200 mM ammonium acetate, pH 8.0, and diluted to 13-15 nM. MP measurements were performed essentially as described previously (Young *et al.*[50]), on a ONE^{MP} instrument (Refeyn Ltd, Oxford) in the case of the *Er*Cry4a WT sample in Fig. 3 and on a TWO^{MP} instrument (Refeyn Ltd, Oxford) in all other cases. Briefly, a borosilicate microscope coverslip (Thorlabs) was cleaned by ultra-sonication for 5 min in a 1:1 mix of ultrapure water and isopropanol followed by another 5 min of ultra-sonication in ultrapure water and then dried under a stream of nitrogen. Cleaned coverslips and Grace Bio-Labs reusable CultureWell^{TM} gaskets (Bio-Labs) were assembled into flow chambers. 10-20 µL of sample were added to the flow chamber and images of a 3.5 × 12.2 µm^2 region of the glass coverslip surface were acquired at 1000 frames s^{-1}. The mass distributions were

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

obtained from contrasts of ratiometric data by calibration using DiscoverMP (Refeyn Ltd, Oxford).

Mass spectrometric analysis of chemically cross-linked peptides

XL-MS was used to probe the interaction surfaces of the linked monomers. Samples for XL-MS were sent to Oxford as described above, with the addition of 2 mM DTT in the shipping buffer to limit higher-order oligomerisation. Furthermore, His-tags were cut leaving only 5 additional amino acids at the N-terminus (GAMGS).

To investigate potential disulphide bonds, protein solutions were incubated with the linkers disuccinimidyl sulphoxide (DSSO)[51] or disuccinimidyl dibutyric acid (DSBU)[52] to cross-link primary amines at the dimer interface using the manufacturer's recommended procedure, as described previously.[53] Cross-linked and non-cross-linked proteins were resolved using SDS-PAGE for analysis of the cross-linkers and direct analysis of the disulphide bonds.

Bands were excised from gels and prepared for mass spectrometry as described by Shevchenko *et al.*[54] After in-gel digestion, peptides were resolubilised in buffer A (H_2O, 0.1% formic acid), and ~5-10 µL of peptide was loaded onto a reverse-phase trap column (Acclaim PepMap 100, 75 µm × 2 cm, NanoViper, C18, 3 µm, 100 Å, ThermoFisher, Waltham, MA, USA) using an Ultimate 3000 autosampler and washed with 40 µL of buffer A at a flow rate of 20 µL min^{-1}. The trapped peptides were then separated by applying a 15 cm reverse-phase analytical column (Acclaim PepMap 100 C18, 3 µm, 150 mm × 0.075 mm) using a 105 min linear gradient from 5% to 55% buffer B (80% acetonitrile, 20% water, 0.1% formic acid) at a flow rate of 300 nL min^{-1}. Long and hydrophobic peptides were washed from the column using 99% buffer B for 15 min at the same flow rate. The separated peptides were electrosprayed in the positive ion mode into an Orbitrap Eclipse Tribrid mass spectrometer (ThermoFisher) operated in data-dependent acquisition mode (3 s cycle time) using the workflow recommended by the DSSO and DSBU methods programmed by the manufacturer. Briefly, precursor scans were collected in the Orbitrap analyser at 60,000 resolving power at *m*/*z* 200 with a mass range of 375-1600 *m*/*z*. Precursors above the intensity threshold of 1.0×10^4 having charge states between $z = 4$ and $z = 8$ were isolated using the quadrupole (0.5 *m*/*z* offset, 1.6 *m*/*z* isolation window) and fragmented in the linear ion trap using collision-induced dissociation (collision energy = 25%). Peptide fragments were analysed using the

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Orbitrap at a resolving power of 30,000 at *m*/*z* 200 with a maximum injection time of 70 ms. Fragment ions with z = 2 to z = 6 spaced by the targeted mass difference of 25.9 Da or 31.9 Da (± 10 ppm) corresponding to peptides cross-linked by DSBU or DSSO, respectively, were subjected to further sequencing in the linear ion trap operated in rapid detection mode (35% collision energy, MS1 isolation of 2.5 *m*/*z* and MS2 isolation of 2.6 *m*/*z*). Additional MS/MS scans for precursors within 10 ppm were dynamically excluded for 30 s following the initial selection. The cross-linking data were analysed using the free software tool MeroX[55] available at www.StavroX.com.

Cysteine exposure measurements

Quantitative determination of thiol groups in solution was achieved essentially as described previously[56] by recording the formation of 5-thio-2-nitrobenzoic acid (TNB) from 5,5′-dithio-bis-(2-nitrobenzoic acid) (DTNB) at a wavelength of 412 nm. Briefly, WT *Er*Cry4a, mutants C317A, C412A and the 5-cysteine mutant were dialysed overnight in 100 mM Tris/HCl at pH 8.0. Next, a fresh 12 mM DTNB solution in 100 mM Tris/HCl pH 8.0 was produced. A calibration curve with 0, 5, 10, 15, 20, 25 and 30 µM L-cysteine and 60 µM DTNB was prepared to calculate the slope of the line (see Fig. S2). After adding DTNB, the samples were incubated for 10 min at room temperature. Afterwards, the absorbance of TNB in the Cry samples was measured at 412 nm. To measure the accessibility of cysteines in the proteins, 5 µM of WT *Er*Cry4a and of the *Er*Cry4a mutants were incubated for 10 min at room temperature with 60 µM DTNB in 100 mM Tris/HCl pH 8.0. The number of accessible cysteines was calculated using the expression, $N_{\mathrm{Cys}} = E / ac$, where E is the total absorbance of the sample, a is the slope of the calibration graph and c is the concentration of protein in the solution. All measurements were done in ambient light conditions.

Computational Methods

Dimer formation through both covalent and non-covalent interactions of surface residues have been investigated for both full-length and truncated *Er*Cry4a structures. All of the dimers studied were classified into families representing groups of structures according to the type of interaction, i.e., covalent (cov) or non-covalent (ncov). The covalent structures were

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

labelled cov(N)^n, where N is the sequence number of the cysteine residue in the monomer subunits that forms the disulphide bond and n labels a particular dimer within the family.

Identifying exposed residues for covalent dimer construction

Covalently linked dimers arise from the formation of disulphide bonds between cysteine residues. Figure 2 summarises the overall workflow for constructing the 16 *Er*Cry4a dimeric structures we have investigated. A previously simulated *Er*Cry4a construct, including the FAD cofactor, was used for an analysis of the solvent exposure of the 11 cysteines in

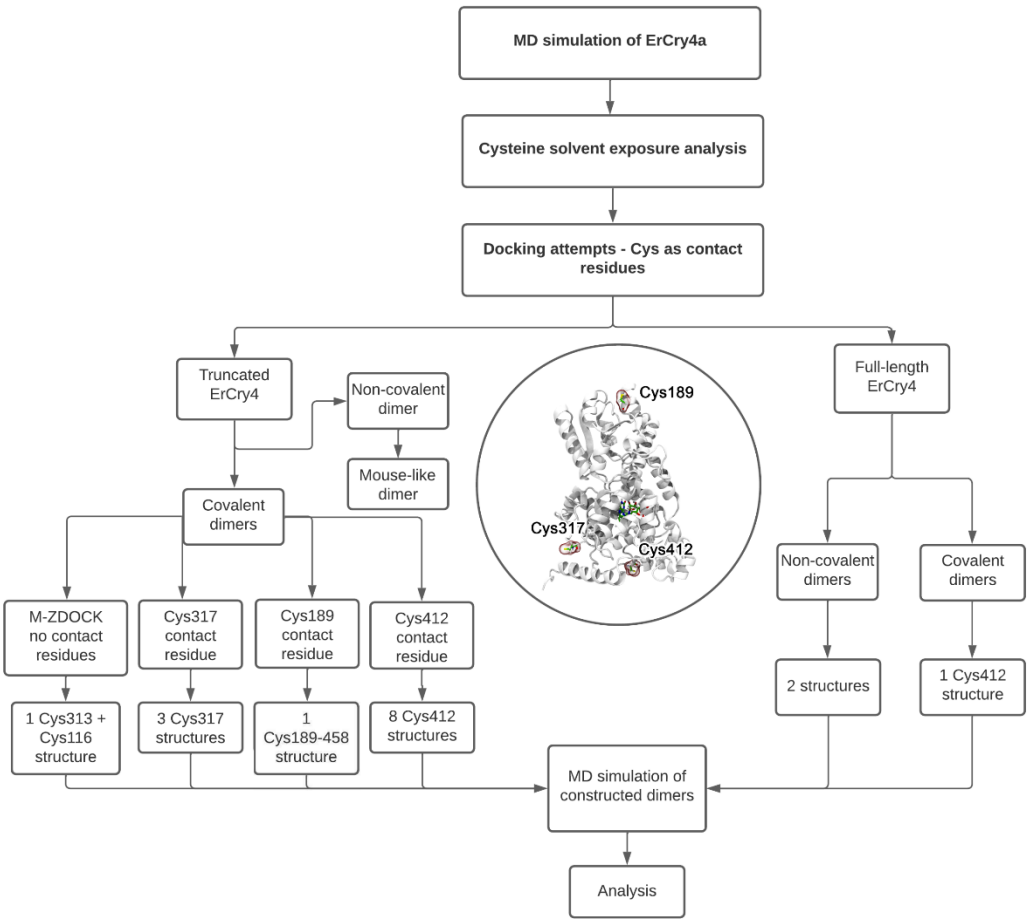

Figure 2. Computational workflow used to produce the 16 dimeric structures of *Er*Cry4a. The inset in the middle highlights the contact cysteine residues (C317, C412 and C189) which were used in constructing the covalent dimers. An overview of all structures obtained during the docking procedure can be found in the Supplementary Material in Tables S10-S18

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

*Er*Cry4a[15]. 50 snapshots, taken from a 200 ns MD simulation, were time-averaged and evaluated using GETAREA2[57]. To compare SASA values for different amino acids, the absolute solvent-accessible surface area for each residue (SASA)[58] was divided by the maximum SASA (MSA) value for the corresponding amino acid (X) in a Gly-X-Gly tripeptide. A similar method was introduced in earlier work[59]. Gly-X-Gly tripeptides were constructed with Pep McConst[60] software and simulated for 1 ns in a waterbox using NAMD. MSA values for the X residues were determined using GETAREA2[57]. The solvent exposure for residue, *i*, was calculated as:

$$S_i = \frac{SASA_i}{MSA_i} \times 100\% . \quad (1)$$

Covalent dimer construction

The ZDOCK docking tool[61–63] was used to create a set of *Er*Cry4a dimers using the seven most solvent-exposed cysteine residues as the docking sites (Table 3). ZDOCK allows one to define a contact residue for dimer construction such that the defined residues in the monomers are placed in close proximity to each other. For the six most solvent-exposed cysteines residues (C68, C73, C116, C189, C317 and C412) together with C179, 49 dimeric structures were constructed, and inter-cysteine distances were analysed to determine the possibility of disulphide bond formation (See Tables S10-S18). Docking of two *Er*Cry4a monomers with the contact residues Cys68 (Table S10), Cys73 (Table S11), Cys116 (Table S12), or Cys179 (Table S13) produced 10 dimeric structures. However, because the cysteine residues were more than 10 Å apart, these structures were not considered in the MD simulations (Table S19). With Cys189 (Table S14) and Cys317 (Table S15) as contact residues, four dimers were found in which one and three structures, respectively, contained the cysteines in close enough contact to warrant proceeding to MD simulations. The disulphide bonds were introduced with NAMD[64] by removing the hydrogen atoms from the cysteine SH groups and defining an abnormally long covalent connection between the sulphur atoms. During the follow-up equilibration, the bond acquired the length of a typical disulphide bond causing a slight rearrangement of atoms at the interface to accommodate the structural change. With Cys412 as the contact residue, ZDOCK yielded three dimeric structures (Table S16) in which

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

the cysteines were somewhat separated, making them unstable in short MD simulations. Since the experimental results (below) suggested that the Cys412 residue was potentially important for dimerisation, we have used snapshots from the short MD simulation to construct and simulate 8 dimeric configurations for the Cys412 family by first imposing an artificial constraint on the Cys412-Cys412 bond to help the residues move closer together, and then removing the constraint.

An additional 10 structures were generated without explicitly specifying a contact residue (Table S17). Furthermore, the M-ZDOCK tool[65] was used to search for different plausible *Er*Cry4a dimer structures by predicting cyclically symmetric multimers based on the structure of the *Er*Cry4a monomer (Table S18). In the case of a dimer, ten symmetric structures were obtained, but only one, referred to as cov^{D}, was analysed further (Table S18). This dimer contained two disulphide bonds: C116-C313 and C313-C116.

Table 3. Summary of *Er*Cry4a dimer structures produced by ZDOCK and M-ZDOCK and the MD simulations. The specified lengths of the simulations apply to each member of the protein family.

Contact Residue	Docked Structures	Simulated Structures	Simulation Time (ns)
Covalent dimeric structures			
68	10	0	–
73	10	0	–
116	10	0	–
179	10	0	–
189	3	1	100
317	3	3	400
412	3	8	150
none	10	1	100
none (M-ZDOCK)	10	1	100
Non-covalent dimeric structures			
none	10	2	400

Non-covalent dimer construction

The computational tool I-TASSER[66–69] was used to generate the intrinsically disordered CTT (residues 498-527) missing from the crystal structure of pigeon Cry4a (PDB code: 6PU0)[28] to obtain a full-length (FL) homology model, *Er*Cry4aFL, with a (confidence) C-score of 0.34. This parameter reports the quality of the predicted model and ranges from –5 to +2,

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

where a higher value signifies a model with a higher confidence (Table S5). The C-score is calculated based on the significance of threading template alignments with subsequent reassembly using replica-exchanged Monte Carlo simulations and the convergence parameters of the structure assembly simulations[66–69].

Structures of possible *Er*Cry4aFL non-covalently bound dimers were generated using ZDOCK[61]. Shape complementarity was taken into consideration, where the algorithm keeps one monomer fixed and rotates and translates the other to find configurations that result in the best fitting score. In the investigation of possible covalent and non-covalent dimers, ZDOCK produced 2000 dimer configurations, of which the top 10 were used for further investigation. These 10 structures were chosen using an energy-based scoring function which includes the potential energy, spatial complementarity, and electric field force between the protein subunits.

Dimeric structure inspired by *Mm*Cry2

Dimer interfaces have a high degree of conservation in evolutionarily related proteins[1]. Even though the crystallographic dimer of *Mm*Cry2 has not been identified as being functionally relevant, it is still the closest protein to *Er*Cry4a for which a dimer crystal structure has been reported (PDB: 6KX8[46]). A dimeric structure inspired by *Mm*Cry2 was thus constructed, in which the monomers were structurally aligned with the *Mm*Cry2 dimer subunits.

Molecular dynamics simulations

All MD simulations were conducted using the NAMD package[64,70] through the VIKING platform[71]. The CHARMM36 force field included CMAP corrections for proteins[72,73] and additional parameterisations for FAD[29,74–76]. Periodic boundary conditions were adopted in all MD simulations and the particle mesh Ewald summation method was employed to evaluate long-range Coulomb interactions. Van der Waals interactions were treated using a smooth cut-off of 12 Å with a switching distance of 10 Å. The simulation temperature was 310 K, controlled with the Langevin thermostat[70]. A constant pressure of 1 atm for equilibrium simulations was obtained using the Langevin piston Nosé–Hoover method[77]. The SHAKE algorithm was used to constrain bonds including hydrogen atoms at their respective equilibrium distances. After 10,000 NAMD[70] minimisation steps, harmonic restraints were

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

initiated in the system and gradually released to achieve an equilibrium structure. Equilibration simulations were conducted with explicit solvent modelled through the TIP3P parameter set[78] with water molecules surrounding the dimers to a distance of 15 Å in all directions. A NaCl salt concentration of 50 mM was assumed in all simulations. After the equilibration, production simulations with temperature control (310 K) within the *NVT* statistical ensembles were performed (Table S19). All MD simulation results were analysed with VMD[79].

Results and discussion

We have studied the dimerisation of *Er*Cry4a using experimental and computational approaches in parallel. Computational results have been used to guide the experiments and *vice versa* at various stages of the investigation.

*Er*Cry4a forms a population of disulphide-linked dimers

Figure 3A shows a native mass spectrum of WT *Er*Cry4a. Two charge-state series consistent with the monomer and dimer can be seen: major peaks correspond to the monomer, and minor peaks in the range m/z = 5000-7000 to the dimer.

To investigate the dimerisation further, WT *Er*Cry4a was studied by mass photometry (Fig. 3B). Two peaks were observed, a major component (60-80 kDa, 86%) from monomers and a minor component (120-160 kDa, 11%) attributed to dimers. The observation of dimeric forms of *Er*Cry4a both by native MS (at micromolar protein concentrations) and MP (at nanomolar concentrations) suggests that the two monomer units are strongly bound and not an artefact of nanoelectrospray ionisation. Non-specific, non-covalent dimers would be expected to dissociate at high dilution. Notably, the relative abundance of dimers and monomers was the same in both the MS and MP experiments despite the large difference in sample concentration, suggesting that the dimers are covalently linked. To test this possibility, we ran SDS-PAGE gels of *Er*Cry4a samples after incubation with various concentrations of DTT reductant (Fig. 3C). Bands from dimers and higher order oligomers, visible under the denaturing conditions of the gels in the absence of DTT, are attenuated by DTT-treatment, consistent with dimer formation via covalent inter-monomer disulphide bonds. The complete gel is shown in Fig. S1.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

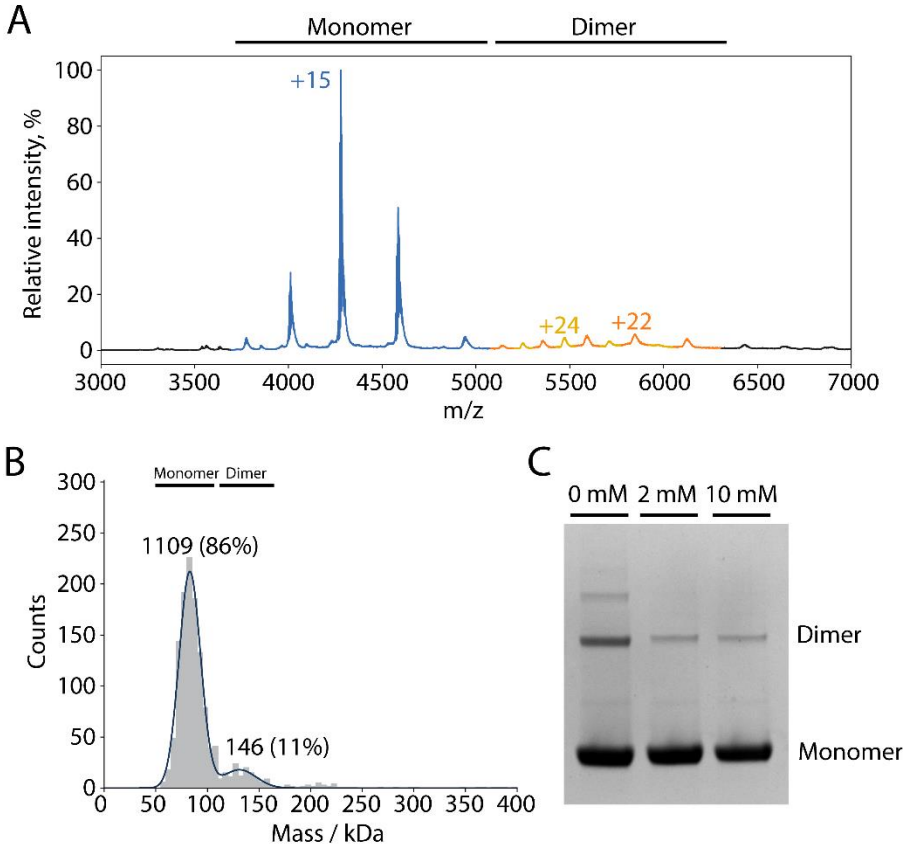

Figure 3. (**A**) Native mass spectrum of *Er*Cry4a. The peaks of the monomer region are indicated in blue (molecular mass 64.197 ± 0.029 kDa) and the peaks of the dimer region in orange and yellow (molecular masses ~123 kDa and ~131 kDa). The expected mass of an *Er*Cry4a WT monomer including FAD and a His-tag is 64.034 kDa. The differences between the detected and the expected masses are due to salt ions and buffer components. The finding of a second dimer mass, lower than the expected mass, could be due to minor truncations and is only observed in some samples. The observation here of stable dimers is consistent with Fig. S13 of Xu et al.[15] The sample used for this measurement contained 10 mM BME during shipment to prevent higher order oligomerisation. (**B**) Mass photometry of 44 nM *Er*Cry4a. The same sample as in (A) was used. The numbers above the peaks are the numbers of counts. (**C**) SDS-PAGE gel of *Er*Cry4a samples without a His-tag (the same as used for XL-MS) that had been incubated with various amounts of DTT.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Cysteine solvent-accessibility calculations

On the basis that the covalent links responsible for dimerisation are most likely to be disulphide bonds, we calculated the solvent-accessible surface area (SASA) of the 11 cysteines in *Er*Cry4a (Fig. 4 and Table S3). Six of these residues have sidechain accessibilities greater than 20%, the other five being less than 7%. The most exposed cysteine, C317, is part of a turn near to the end of the Trp-tetrad. The next four most exposed are all located on or close to an α-helix.

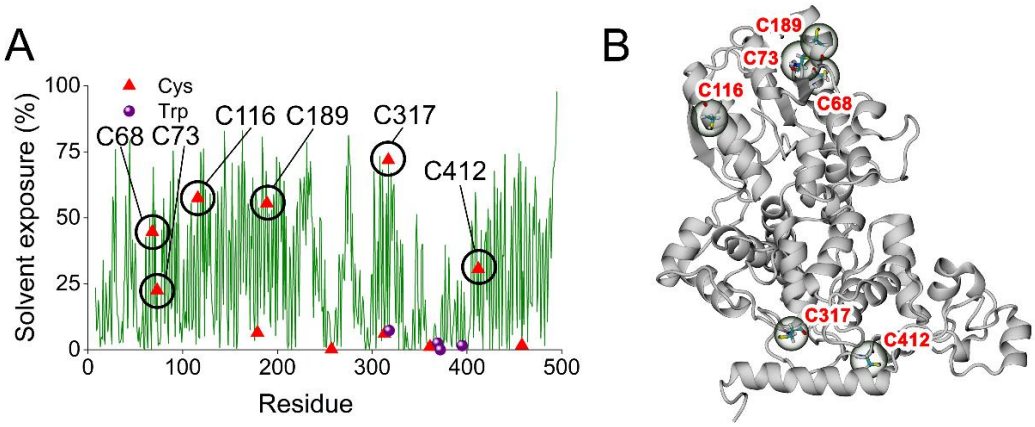

Figure 4. (**A**) Solvent exposures of residues in monomeric WT *Er*Cry4a with cysteines (red triangles) indicated. Also shown (purple spheres) are the components of the Trp-tetrad: (W318, W369, W372, W395). (**B**) *Er*Cry4a secondary structural elements with the six most exposed cysteines indicated.

Dimerisation of cysteine mutants

M-ZDOCK was used to identify potential dimers by docking pairs of monomers without specifying a contact residue (see below). Using a selection criterion that the sulphur atoms of the cysteines should be closer than ~6 Å, three residues, C116, C313, and C317, were identified as possible components of intermolecular disulphides (Table S18). C116 and C317 have the highest estimated solvent exposure of all 11 cysteines, while C313 is largely buried (Fig. 4 and Table S3). Four mutants of *Er*Cry4a were expressed and purified with one, two or all three of these cysteines replaced by alanines: C317A, C116A+C313A, C116A+C317A,

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

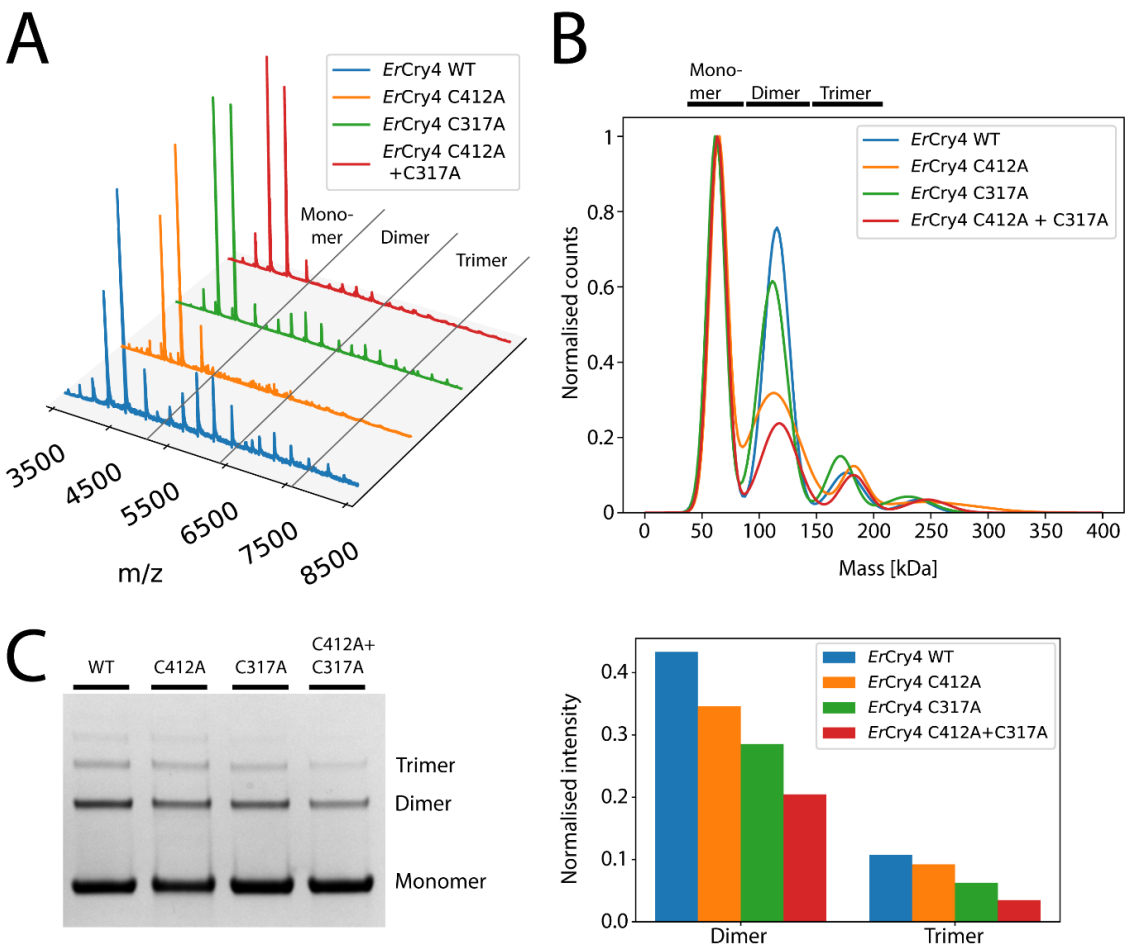

Figure 5. (**A**) Native mass spectra of WT *Er*Cry4a and three cysteine mutants. All four proteins are mostly monomeric but also show charge-state envelopes from dimers and trimers. (**B**) Mass photometry results for the same proteins at concentrations between 13-15 nM displayed as Gaussian-fitted normalised counts. The four traces have been scaled so that the monomer peaks have the same height. (**C**) Denaturing SDS-PAGE gel for the same four proteins (full SDS-PAGE gels can be found in Fig. S1C) and the normalised intensities of the protein areas.

and C116A+C313A+C317A. As judged by SDS-PAGE (Fig. S1B), all four mutants oligomerise, implying that if cysteines C116, C313 and C317 are involved in dimer formation, they are not the only ones.

Two further *Er*Cry4a mutants were investigated: C412A (chosen based on docking followed by MD simulations and on chemical cross-linking followed by mass spectrometric analysis, see below) and C68A+C73A+C116A+C189A+C317A (with five of the six

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

cysteines with the largest calculated SASAs replaced by alanines). Both proteins oligomerised showing that these mutations, like the four above, might reduce the ability of *Er*Cry4a to dimerise but fail to remove it completely.

The *Er*Cry4a mutants C412A, C317A, and C412A+C317A were subsequently chosen for further comparative investigation. As seen in Fig. 5, all three have a lower degree of dimerisation than the WT protein, and the double mutant dimerises less than either of the single mutants in both MP and on the gels. In the MP measurements (Fig. 5B), the degree of dimer formation is in the order WT > C317A > C412A > C412A+C317A. The low concentrations used in this experiment (13-15 nM) suggest that this dimerisation is largely covalent and involves both cysteines. The reversed order of C317A and C412A in Fig. 5C perhaps indicates a different tendency for non-covalent and covalent association. The presence of trimers in the denaturing gel necessitates that two or more disulphide links are required for their formation.

Cysteine accessibility measurements

The average number of accessible cysteines in WT and mutant forms of *Er*Cry4a was determined by measuring the absorption of the product formed by treating the protein with Ellman's reagent (DTNB). The assay finds approximately five accessible cysteines for the wild-type protein, four for each of the C317A and C412A mutants and three for the C116A+C317A double mutant (Table 4). More detail on these absorbance measurements is given in Table S1. Assuming that each cysteine is either completely unreactive towards DTNB due to structural hindrance or reacts stoichiometrically, these results suggest that all three of C317, C412 and C116 are significantly exposed on the surface of the protein and in principle available for the formation of intermolecular disulphide bonds. However, as a note of caution, a cysteine able to react with DTNB (a small molecule) is not necessarily accessible enough for disulphide formation between monomers.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 4. TNB absorbance measurements (at 412 nm) and average number of accessible cysteines in wild-type and mutant forms of *Er*Cry4a

	A_{412}	Accessible Cysteines
WT	0.30 ± 0.01	4.78 ± 0.13
C317A	0.27 ± 0.01	4.23 ± 0.10
C412A	0.27 ± 0.00	4.16 ± 0.05
C116A+C317A	0.17 ± 0.01	3.30 ± 0.11

Cross-linking mass spectrometry

XL-MS was used to investigate the interface between monomer units in *Er*Cry4a dimers by creating artificial chemical cross-links that can be cleaved by MS so as to reveal which parts of the monomers come into close contact in the dimers. The bifunctional compounds DSSO and DSBU cross-link the exposed primary amine groups of lysine residues, via bis(N-hydroxysuccinimide) ester groups. MS analysis of peptides formed by enzymatic digestion of the cross-linked dimers identifies pairs of lysines with α-carbons separated by up to ~2.7 nm thereby constraining the location and geometry of the protein-protein interface.

Five lysines were identified to form cross-links. Based on whether they were found in the monomeric or dimeric fractions of the protein, two links were assigned to be intramolecular and three to be intermolecular, as displayed in Fig. 6A. Four disulphide bonds were identified (C361-C458, C412-C361, C412-C412, and C412-C458), three of which involve C412. Of these, only C412-C412 is unequivocally intermolecular: the difficulty of distinguishing the two possibilities in a homodimer means that the other three could be intermolecular or intramolecular or both. The intermolecular K152-K152 link is consistent with a C412-C412 disulphide bond because K152 and C412 are on the same side of the protein (Fig. 6B). The links from K234 to K429 and from K152 to K507 were also found in the monomer protein fraction suggesting that they are more likely to be intra-monomer. For more details see Table S2. Overall, these results point to C412 as a likely component of disulphide links between monomers.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

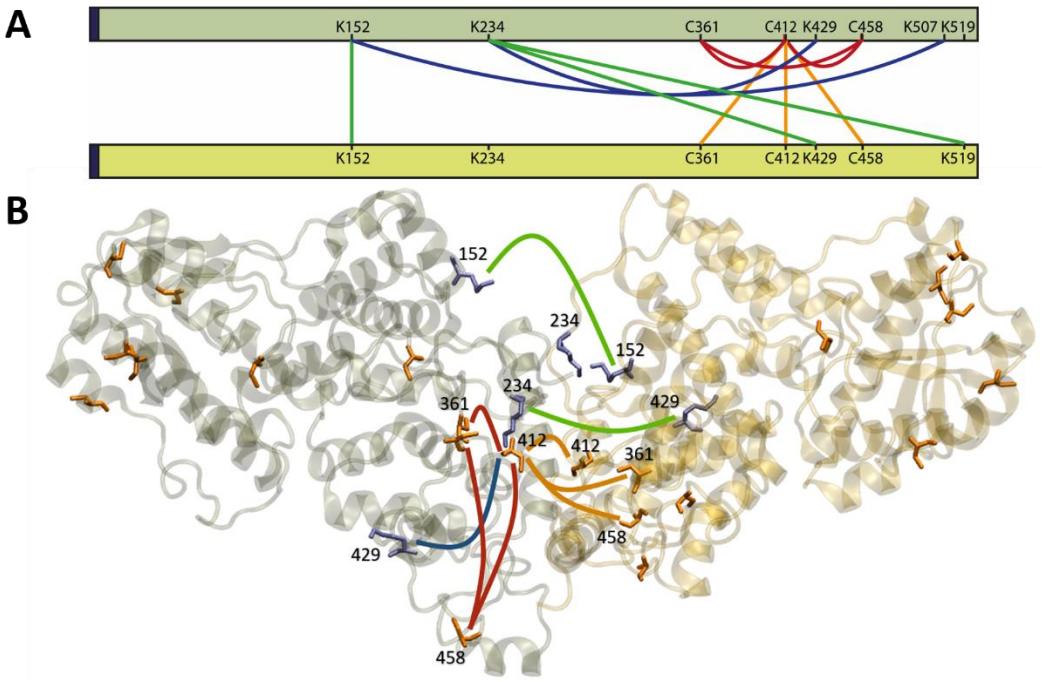

Figure 6. (**A**) Network plot of cross-linking sites found for *Er*Cry4a. The orange and red lines are potential inter- and intramonomer disulphide bonds, respectively. The inter- and intramonomer cross-links formed by DSBU and DSSO are shown in green and blue, respectively. The dark blue region at the N-terminus (left hand side) corresponds to the five additional amino acids that remained after His-tag cleavage. (**B**) Graphical representation of a possible coordination of two *Er*Cry4a monomers. The cross-links are drawn with the same colour code as in (A). Cysteine residues are shown in orange and lysine residues in purple. Only residues 1-495 are displayed.

Computationally reconstructed *Er*Cry4a dimer structures

As described in the Methods section (Fig. 2), a large number of covalently and non-covalently bound dimers were constructed using ZDOCK and M-ZDOCK. A total of 49 docked structures were generated for a truncated form of *Er*Cry4a (Table 3) using the seven most exposed cysteines as contact residues. Another ten were obtained without specifying which cysteines should be near one another in the docked structure. Twelve of these structures were selected for MD simulations, based on the separations of the cysteine residues (eight with a C412-C412 disulphide, three with C317-C317, one with C189-C485). To these were added:

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

a non-covalently bound structure derived from the crystallographic dimer of mouse Cry2, a dimer with two disulphide bridges, and two structures for the full-length form of *Er*Cry4a, both non-covalent.

Of the 16 dimer structures chosen for MD simulation (see Methods section and Supplementary Material for more details), seven were selected (Table 5) on the basis of their monomer-monomer interaction energies, E_{tot} : two structures with C412-C412 disulphides (cov^{412A} and cov^{412B}), two with C317-C317 bonds (cov^{317A} and cov^{317B}), two non-covalent ($ncov^{A}$ and $ncov^{M}$) and the one with two disulphides (cov^{D}, C116-C313 and C313-C116). $ncov^{M}$ is the structure based on the crystallographic dimer of *Mm*Cry2.

Various other properties of these structures, derived from the MD simulations, are presented in Table 5 and described below. Of the seven dimers, $ncov^{A}$ and cov^{412A} have the largest internal potential energies (E_{tot}) and are therefore the most stable. The interaction between the monomers in $ncov^{A}$ is probably stronger because the disulphide bond in cov^{412A} (and the other covalent structures) constrains the monomers at the dimer interface preventing them finding more favourable binding conformations. Note that E_{tot} comprises the non-covalent van der Waals and electrostatic interactions between the monomers and does not include the ca. 60 kcal mol^{-1} expected for a disulphide bond[80]. The other dimers featured in Table 5, especially cov^{317A}, cov^{D}, and cov^{317B}, have smaller interaction energies and are therefore less likely to be formed.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 5. The seven simulated and analysed *Er*Cry4a dimers with the largest non-bonded monomer-monomer interaction energies, E_{tot}. Results for other dimers are given in the Supplementary Material. The various parameters are discussed in the text.

Dimer	$-E_{tot}$ / kcal mol^{-1}	R_g /Å	RMSD/ Å	RMSF/ Å	A_{IS} / Å^2	Hydrogen bonds	Salt bridges
ncov^{A}	857± 85	35.2±0.3	3.5±0.4	2.2±1.0	2762±292	106	29
cov^{412A}	526±83	41.2±0.2	2.5±0.4	5.7±1.3	1840±115	31	13
ncov^{M}	505±163	33.0±0.3	3.4±1.4	2.3±1.0	2442±284	111	12
cov^{412B}	437±127	41.4±0.7	6.2±2.8	5.4±1.5	1040±310	49	12
cov^{317A}	186±103	38.3±0.3	4.0±0.6	1.8±0.9	1082±220	64	10
cov^{D}	178±44	33.0±0.2	2.5±0.3	1.3±0.7	2140±330	47	3
cov^{317B}	171±73	38.5±0.2	3.8±0.8	15.6±4.1	1070±110	59	12

The stronger binding of the two non-covalent dimers is reflected in their smaller radius of gyration, R_g, which is a measure of the compactness of the structure. The cov^{D} dimer also has a relatively small R_g presumably because of the constraining effects of its two disulphide bonds.

Another indication of protein stability is provided by the average root-mean-squared deviation (RMSD) of the backbone atom positions from those in a reference structure immediately prior to the production simulation[81]. Smaller values of this parameter correspond to dimers in which the initial docked structure did not alter significantly throughout the simulation. Apart from cov^{412B} (6.2 Å), all the dimers in Table 5 have RMSDs less than 4 Å. Dimers cov^{412A} and cov^{D} (RMSD ≈ 2.5 Å) are the most stable as judged by this measure.

A parameter that quantifies the degree of internal mobility of the dimer is the root-mean-square amplitude of the structural fluctuations (RMSF) during the course of an MD simulation. RMSF values were calculated by comparing the positions of each carbon atom at every simulation step to its position in the average structure. With the exception of cov^{317B}, which is much less rigid than the others, all the structures in Table 5 have comparable values of this parameter. Certain regions of the monomers are more flexible than others, in particular the C-terminal extension (residues 470-495) and the phosphate-binding loop (residues 231-248). For the most part, this flexibility seems to be retained on dimerisation.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

A further measure of the interaction between the monomer units of a dimer is the interaction surface area, A_{IS} (Table 5), defined as the difference between the solvent accessible surface areas of the separate monomers and that of the dimer. Larger values of A_{IS} should correlate with stronger binding energies ($-E_{tot}$) as seems to be broadly the case for the dimers in Table 5. Two additional parameters are the number of hydrogen bonds and the number of salt bridges that link the two components of the dimer. Only salt bridges present in more than 10% of the MD frames were counted. All values have been averaged over the duration of the corresponding MD simulations. The two non-covalent structures have the largest interaction surface areas and many more hydrogen bonds than the covalently bound dimers.

Figures 7A and B show the dimer interfaces of cov^{412A} and $ncov^{A}$. Consistent with the data in Table 5, the non-covalent dimer has a more extensive interface.

As described in the Methods section, none of the docked dimers obtained using C68, C73, C116 and C179 as the contact residue were deemed worth simulating by MD. These four residues, together the four with the smallest solvent accessibilities (C257, C313, C361, C458) appear to have a low propensity to dimerise via the formation of a single disulphide bond. All of the C189-linked dimers which satisfied our conditions for MD simulation had smaller binding energies than the seven in Table 5 and therefore appear less likely to be formed. Of the covalently bound dimers, this leaves C317-C317, C412-C412 and the double-disulphide dimer, cov^{D}, which involves C116 and C313. Based on the data in Table 5, C412 seems most likely to form covalent links between *Er*Cry4a monomers. The other main conclusion that can be drawn from these studies is that the possibility of non-covalently bound dimers cannot be dismissed.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

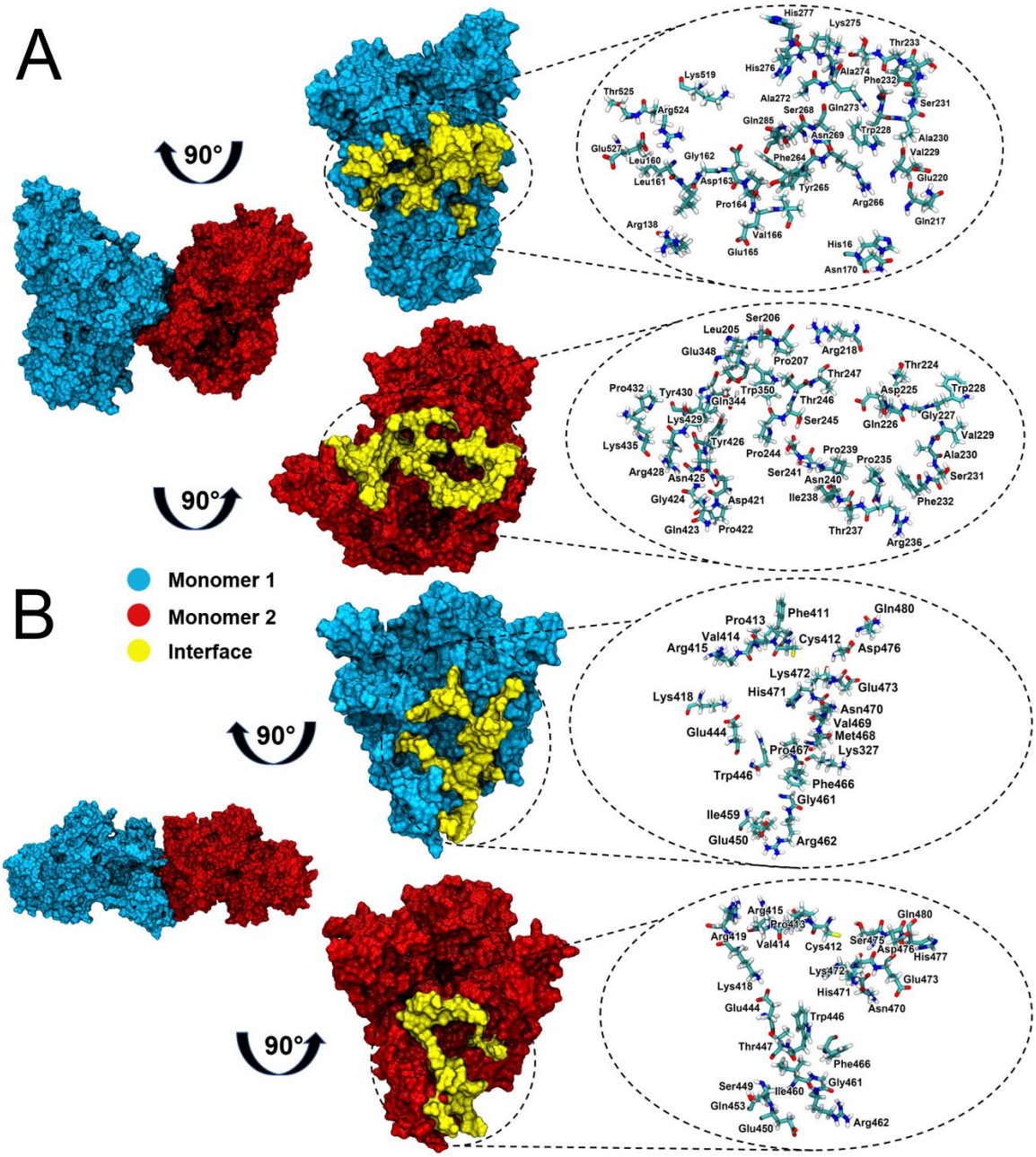

Figure 7. The dimer interfaces of (**A**) ncov^{A} and (**B**) cov^{412A}. Left: dimer structures. Middle: 90°-rotated monomers showing the interacting residues (yellow). Right: representations of the contact residues in the interface.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Conclusions

The experiments described here show that *Er*Cry4a readily forms dimers and that they are covalently linked by disulphide bonds. Several cysteines seem to be involved, C317 and C412 being the most likely. Computational modelling supports this conclusion, provides insights into the relative orientation of the monomer units, and suggests that non-covalently bound dimers may also be relatively stable. The relevance of these findings to the proposed role of Cry4a in avian magnetoreception is unclear.

That *Er*Cry4a dimers are not readily disrupted in a reducing environment (10 mM DTT, Fig. 3) suggests that they could persist *in vivo*. Despite the reducing environment of the cell, there are well-documented examples of biologically significant intra-cellular disulphide-linked proteins involved in redox processes[82–84]. Moreover, little is known about the redox conditions in the avian photoreceptor cells that are thought to contain the magnetoreceptors. It therefore seems premature to exclude the possibility that *Er*Cry4a dimerisation could have a biological function.

If *Er*Cry4a does have a role in magnetic sensing and/or signalling, then an immediate question is why the protein has evolved to stabilise inter-monomer disulphide bonds to the extent that only a proportion of the proteins dimerise under the conditions of the experiments reported here. One explanation is that these dimers are simply aggregates of unfolded monomers. However, we think this is unlikely, as the *Er*Cry4a dimers bind FAD and have native MS charge-state distributions that are similar to those of the monomers suggesting that the protein remains correctly folded and therefore potentially functional in the dimeric state. Partially unfolded proteins typically display higher charge states than the native states. An alternative explanation is that a monomer-dimer equilibrium could play a regulatory role in magnetic sensing or signal transduction. Activation of plant cryptochromes 1 and 2 leads to homo-oligomerisation[3,5,40,41,85]. Additionally, the formation of a disulphide bond influences the interaction between mouse cryptochrome 1 and the circadian protein Period (PER) and depends on the metabolic/oxidative state inside the cell[86]. Taken together, these examples show that photoactivation of some cryptochromes causes oligomerisation and that cryptochrome's affinity for other proteins could rely on disulphide bond formation.

An observation reported by Xu et al.[15] (in their Fig. S5) may be relevant here. Samples of *Er*Cry4a that had been purified without adding 10 mM BME to prevent dimerisation showed clear evidence for the formation of relatively long-lived (nanosecond to microsecond) photo-excited singlet (^{S}FAD*) and triplet (^{T}FAD*) states of FAD in addition to the normal photo-induced electron transfer reactions that generate radical pairs. It was speculated that subtle differences in protein conformation, induced by dimerisation, might inhibit electron transfer from the tryptophan tetrad to ^{S}FAD*, thereby

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

stabilising ^{S}FAD* and allowing formation of ^{T}FAD* by intersystem crossing. Samples of *Er*Cry4a purified with 10 mM BME, and therefore more likely to be monomeric, showed much weaker signals from ^{S}FAD* and ^{T}FAD*.

While recognising that the experiments described here have not been performed in cells, they do show it is quite possible that dimerisation could have an impact on the *in vivo* function of avian cryptochromes and provide an incentive for further work in this area.

Supporting Information

Supplementary figures S1-S14 and Tables S1-S20 generated in the present study. Additional supplementary data and protocols used in the present investigations are provided.

Acknowledgments

The authors would like to thank the Volkswagen Foundation (Lichtenberg Professorship to IAS), the Deutsche Forschungsgemeinschaft (DFG; GRK1885: Molecular Basis of Sensory Biology and SFB 1372: Magnetoreception and Navigation in Vertebrates), the Ministry for Science and Culture of Lower Saxony (Simulations Meet Experiments on the Nanoscale: Opening up the Quantum World to Artificial Intelligence (SMART) and Dynamik auf der Nanoskala: Von kohärenten Elementarprozessen zur Funktionalität (DyNano)), the Japan Science and Technology Agency (JST) PRESTO (Quantum Bio) Grant No. JPMJPR19G1 (LMA), the European Research Council (under the European Union's Horizon 2020 research and innovation programme, Grant Agreement No. 810002, Synergy Grant: *QuantumBirds*), and the Office of Naval Research Global, award no. N62909-19-1-2045. Computational resources for the simulations were provided by the CARL Cluster at the Carl-von-Ossietzky University, Oldenburg, supported by the DFG and the Ministry for Science and Culture of Lower Saxony. The authors gratefully acknowledge the computing time granted by the Resource Allocation Board and provided on the supercomputer Lise and Emmy at NHR@ZIB and NHR@Göttingen as part of the NHR infrastructure. The calculations for this research were conducted with computing resources under the project nip00058. TJE was supported by funds from the Royal Society under a Royal Society Newton International Fellowship. ASG is grateful to Mark Pepys who provided CRP for the mass spectrometric investigation and to Josh Bishop who helped with the mass photometry measurements and data analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

References

(1) Matthews, J. M. *Protein Dimerization and Oligomerization in Biology*; Springer Science+Business Media, LLC: Sydney, Australia, 2012.

(2) Heldin, C.-H. Dimerization of Cell Surface Receptors in Signal Transduction. *Cell* **1995**, *80*, 213–223.

(3) Palayam, M.; Ganapathy, J.; Guercio, A. M.; Tal, L.; Deck, S. L.; Shabek, N. Structural Insights into Photoactivation of Plant Cryptochrome-2. *Commun. Biol.* **2021**, *4*, 1–11.

(4) Poehn, B.; Krishnan, S.; Zurl, M.; Coric, A.; Rokvic, D.; Orel, L.; Raible, F.; Wolf, E.; Tessmar-Raible, K. A Cryptochrome Adopts Distinct Moon- and Sunlight States and Functions as Moonlight Interpreter in Monthly Oscillator Entrainment. *bioRxiv* **2021**, 1–12.

(5) Shao, K.; Zhang, X.; Li, X.; Hao, Y.; Huang, X.; Ma, M.; Zhang, M.; Yu, F.; Liu, H.; Zhang, P. The Oligomeric Structures of Plant Cryptochromes. *Nat. Struct. Mol. Biol.* **2020**, *27*, 480–488.

(6) Yu, X.; Liu, H.; Klejnot, J.; Lin, C. The Cryptochrome Blue Light Receptors. In *The Arabidopsis book*; 2010; Vol. 83, pp 1–27.

(7) Chaves, I.; Pokorny, R.; Byrdin, M.; Hoang, N.; Ritz, T.; Brettel, K.; Essen, L. O.; Van Der Horst, G. T. J.; Batschauer, A.; Ahmad, M. The Cryptochromes: Blue Light Photoreceptors in Plants and Animals. *Annu. Rev. Plant Biol.* **2011**, *62*, 335–364.

(8) Brudler, R.; Hitomi, K.; Daiyasu, H.; Toh, H.; Kucho, K. I.; Ishiura, M.; Kanehisa, M.; Roberts, V. A.; Todo, T.; Tainer, J. A.; Getzoff, E. D. Identification of a New Cryptochrome Class: Structure, Function, and Evolution. *Mol. Cell* **2003**, *11*, 59–67.

(9) Navarro, E.; Niemann, N.; Kock, D.; Dadaeva, T.; Gutiérrez, G.; Engelsdorf, T.; Kiontke, S.; Corrochano, L. M.; Batschauer, A.; Garre, V. The DASH-Type Cryptochrome from the Fungus *Mucor Circinelloides* Is a Canonical CPD-Photolyase. *Curr. Biol.* **2020**, *30*, 4483-4490.e4.

(10) Michael, A. K.; Fribourgh, J. L.; Van Gelder, R. N.; Partch, C. L. Animal Cryptochromes: Divergent Roles in Light Perception, Circadian Timekeeping and Beyond. *Photochem. Photobiol.* **2017**, *93*, 128–140.

(11) Cashmore, A. R.; Jarillo, J. A.; Wu, Y. J.; Liu, D. Cryptochromes: Blue Light Receptors for Plants and Animals. *Science (80-.).* **1999**, *284*, 760–765.

(12) Ozturk, N. Phylogenetic and Functional Classification of the Photolyase/Cryptochrome Family. *Photochem. Photobiol.* **2017**, *93*, 104–111.

(13) Malhotra, K.; Sang-Tae, K.; Batschauer, A.; Dawut, L.; Sancar, A. Putative Blue-Light Photoreceptors from *Arabidopsis Thaliana* and *Sinapis Alba* with a High Degree of Sequence Homology to DNA Photolyase Contain the Two Photolyase Cofactors but Lack DNA Repair Activity. *Biochemistry* **1995**, *34*, 6892–6899.

(14) Parico, G. C. G.; Partch, C. L. The Tail of Cryptochromes: An Intrinsically Disordered Cog within the Mammalian Circadian Clock. *Cell Commun. Signal.* **2020**, *18*, 1–9.

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

(15) Xu, J.; Jarocha, L. E.; Zollitsch, T.; Konowalczyk, M.; Henbest, K. B.; Richert, S.; Golesworthy, M. J.; Schmidt, J.; Déjean, V.; Sowood, D. J. C.; Bassetto, M.; Luo, J.; Walton, J. R.; Fleming, J.; Wei, Y.; Pitcher, T. L.; Moise, G.; Herrmann, M.; Yin, H.; Wu, H.; Bartölke, R.; Käsehagen, S. J.; Horst, S.; Dautaj, G.; Murton, P. D. F.; Gehrckens, A. S.; Chelliah, Y.; Takahashi, J. S.; Koch, K.-W.; Weber, S.; Solov'yov, I. A.; Xie, C.; Mackenzie, S. R.; Timmel, C. R.; Mouritsen, H.; Hore, P. J. Magnetic Sensitivity of Cryptochrome 4 from a Migratory Songbird. *Nature* **2021**, *594*, 535–540.

(16) Ritz, T.; Adem, S.; Schulten, K. A Model for Photoreceptor-Based Magnetoreception in Birds. *Biophys. J.* **2000**, *78*, 707–718.

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

(18) Hore, P. J.; Mouritsen, H. The Radical-Pair Mechanism of Magnetoreception. *Annu. Rev. Biophys.* **2016**, *45*, 299–344.

(19) Wong, S. Y.; Solov'yov, I. A.; Hore, P. J.; Kattnig, D. R. Nuclear Polarization Effects in Cryptochrome-Based Magnetoreception. *J. Chem. Phys.* **2021**, *154*, 1–10.

(20) Maeda, K.; Robinson, A. J.; Henbest, K. B.; Hogben, H. J.; Biskup, T.; Ahmad, M. Magnetically Sensitive Light-Induced Reactions in Cryptochrome Are Consistent with Its Proposed Role as a Magnetoreceptor. *PNAS* **2012**, *109*, 4774–4779.

(21) Mouritsen, H.; Janssen-Bienhold, U.; Liedvogel, M.; Feenders, G.; Stalleicken, J.; Dirks, P.; Weiler, R. Cryptochromes and Neuronal-Activity Markers Colocalize in the Retina of Migratory Birds during Magnetic Orientation. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14294–14299.

(22) Möller, A.; Sagasser, S.; Wiltschko, W.; Schierwater, B. Retinal Cryptochrome in a Migratory Passerine Bird: A Possible Transducer for the Avian Magnetic Compass. *Naturwissenschaften* **2004**, *91*, 585–588.

(23) Nießner, C.; Denzau, S.; Gross, J. C.; Peichl, L.; Bischof, H. J.; Fleissner, G.; Wiltschko, W.; Wiltschko, R. Avian Ultraviolet/Violet Cones Identified as Probable Magnetoreceptors. *PLoS One* **2011**, *6*, 1–8.

(24) Nießner, C.; Gross, J. C.; Denzau, S.; Peichl, L.; Fleissner, G.; Wiltschko, W.; Wiltschko, R. Seasonally Changing Cryptochrome 1b Expression in the Retinal Ganglion Cells of a Migrating Passerine Bird. *PLoS One* **2016**, *11*, 1–10.

(25) Bolte, P.; Bleibaum, F.; Einwich, A.; Günther, A.; Liedvogel, M.; Heyers, D.; Depping, A.; Wöhlbrand, L.; Rabus, R.; Janssen-bienhold, U.; Mouritsen, H. Localisation of the Putative Magnetoreceptive Protein Cryptochrome 1b in the Retinae of Migratory Birds and Homing Pigeons. *PLoS One* **2016**, *11*, 1–17.

(26) Bolte, P.; Einwich, A.; Seth, P. K.; Chetverikova, R.; Heyers, D.; Wojahn, I.; Janssen-Bienhold, U.; Feederle, R.; Hore, P.; Dedek, K.; Mouritsen, H. Cryptochrome 1a Localisation in Light- and Dark-Adapted Retinae of Several Migratory and Non-Migratory Bird Species: No Signs of Light-Dependent Activation. *Ethol. Ecol. Evol.* **2021**, *33*, 248–272.

(27) Einwich, A.; Seth, P. K.; Bartölke, R.; Bolte, P.; Feederle, R.; Dedek, K.; Mouritsen,

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

H. Localisation of Cryptochrome 2 in the Avian Retina. *J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol.* **2022**, *208*, 69–81.

(28) Zoltowski, B. D.; Chelliah, Y.; Wickramaratne, A.; Jarocha, L.; Karki, N.; Xu, W.; Mouritsen, H.; Hore, P. J.; Hibbs, R. E.; Green, C. B.; Takahashi, J. S. Chemical and Structural Analysis of a Photoactive Vertebrate Cryptochrome from Pigeon. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 19449–19457.

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

(30) Ozturk, N.; Selby, C. P.; Song, S. H.; Ye, R.; Tan, C.; Kao, Y. T.; Zhong, D.; Sancar, A. Comparative Photochemistry of Animal Type 1 and Type 4 Cryptochromes. *Biochemistry* **2009**, *48*, 8585–8593.

(31) Hochstoeger, T.; Said, T. Al; Maestre, D.; Walter, F.; Vilceanu, A.; Pedron, M.; Cushion, T. D.; Snider, W.; Nimpf, S.; Nordmann, G. C.; Landler, L.; Edelman, N.; Kruppa, L.; Dürnberger, G.; Mechtler, K.; Schuechner, S.; Ogris, E.; Malkemper, E. P.; Weber, S.; Schleicher, E.; Keays, D. A. The Biophysical, Molecular, and Anatomical Landscape of Pigeon CRY4: A Candidate Light-Based Quantal Magnetosensor. *Sci. Adv.* **2020**, *6*.

(32) Chetverikova, R.; Dautaj, G.; Schwigon, L.; Dedek, K.; Mouritsen, H. Double Cones in the Avian Retina Form an Oriented Mosaic Which Might Facilitate Magnetoreception and/or Polarized Light Sensing. *J. R. Soc. Interface* **2022**, *19*.

(33) Balay, S. D.; Hochstoeger, T.; Vilceanu, A.; Malkemper, E. P.; Snider, W.; Dürnberger, G.; Mechtler, K.; Schuechner, S.; Ogris, E.; Nordmann, G. C.; Ushakova, L.; Nimpf, S.; Keays, D. A. The Expression, Localisation and Interactome of Pigeon CRY2. *Sci. Rep.* **2021**, *11*, 1–13.

(34) Einwich, A.; Dedek, K.; Seth, P. K.; Laubinger, S.; Mouritsen, H. A Novel Isoform of Cryptochrome 4 (Cry4b) Is Expressed in the Retina of a Night-Migratory Songbird. *Sci. Reps.* **2020**, *4*, 15794.

(35) Kattnig, D. R.; Nielsen, C.; Solov'yov, I. A. Molecular Dynamics Simulations Disclose Early Stages of the Photo-Activation of Cryptochrome 4. *New J. Phys.* **2018**, *20*, 1–18.

(36) Schuhmann, F.; Kattnig, D. R.; Solov'yov, I. A. Exploring Post-Activation Conformational Changes in Pigeon Cryptochrome 4. *J. Phys. Chem. B* **2021**, *372*, 9652–9659.

(37) Chandrasekaran, S.; Schneps, C. M.; Dunleavy, R.; Lin, C.; DeOliveira, C. C.; Ganguly, A.; Crane, B. R. Tuning Flavin Environment to Detect and Control Light-Induced Conformational Switching in *Drosophila* Cryptochrome. *Commun. Biol.* **2021**, *4*, 1–12.

(38) Watari, R.; Yamaguchi, C.; Zemba, W.; Kubo, Y.; Okano, K.; Okano, T. Light-Dependent Structural Change of Chicken Retinal Cryptochrome 4. *J. Biol. Chem.* **2012**, *287*, 42634–42641.

(39) Kondoh, M.; Shiraishi, C.; Müller, P.; Ahmad, M.; Hitomi, K.; Getzoff, E. D.; Terazima, M. Light-Induced Conformational Changes in Full-Length *Arabidopsis Thaliana*

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Cryptochrome. *J. Mol. Biol.* **2011**, *413*, 128–137.

(40) Rosenfeldt, G.; Viana, R. M.; Mootz, H. D.; Von Arnim, A. G.; Batschauer, A. Chemically Induced and Light-Independent Cryptochrome Photoreceptor Activation. *Mol. Plant* **2008**, *1*, 4–14.

(41) Sang, Y.; Li, Q. H.; Rubio, V.; Zhang, Y. C.; Mao, J.; Deng, X. W.; Yang, H. Q. N-Terminal Domain-Mediated Homodimerization Is Required for Photoreceptor Activity of *Arabidopsis* Cryptochrome 1. *Plant Cell* **2005**, *17*, 1569–1584.

(42) Park, H.; Kim, N. Y.; Lee, S.; Kim, N.; Kim, J.; Heo, W. Do. Optogenetic Protein Clustering through Fluorescent Protein Tagging and Extension of CRY2. *Nat. Commun.* **2017**, *8*, 1–7.

(43) Wu, H.; Scholten, A.; Einwich, A.; Mouritsen, H.; Koch, K. W. Protein-Protein Interaction of the Putative Magnetoreceptor Cryptochrome 4 Expressed in the Avian Retina. *Sci. Rep.* **2020**, *10*, 1–13.

(44) Görtemaker, K.; Yee, C.; Bartölke, R.; Behrmann, H.; Voß, J.; Schmidt, J. Direct Interaction of Avian Cryptochrome 4 with a Cone Specific G- Protein. *Cells* **2022**, *4*, 2043.

(45) Levy, C.; Zoltowski, B. D.; Jones, A. R.; Vaidya, A. T.; Top, D.; Widom, J.; Young, M. W.; Scrutton, N. S.; Crane, B. R.; Leys, D. Updated Structure of *Drosophila* Cryptochrome. *Nature* **2013**, *495*, E3–E4.

(46) Miller, S.; Son, Y. L.; Aikawa, Y.; Makino, E.; Nagai, Y.; Srivastava, A.; Oshima, T.; Sugiyama, A.; Hara, A.; Abe, K.; Hirata, K.; Oishi, S.; Hagihara, S.; Sato, A.; Tama, F.; Itami, K.; Kay, S. A.; Hatori, M.; Hirota, T. Isoform-Selective Regulation of Mammalian Cryptochromes. *Nat. Chem. Biol.* **2020**, *16*, 676–685.

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

(48) Kondrat, D. L. F.; Struwe, W. B.; Benesch, J. L. P. Structural Proteomics: High-Throughput Methods: Second Edition. In *Structural Proteomics: High-Throughput Methods: Second Edition*; Owens, R. J., Ed.; Springer Science \& Business Media: New York, 2015; Vol. 1261, pp 1–371.

(49) Lazar, I. J.; Lazar, I. S. *GelAnalyzer 19.1*. www.gelanalyzer.com.

(50) Young, G.; Hundt, N.; Cole, D.; Fineberg, A.; Andrecka, J.; Tyler, A.; Olerinyova, A.; Ansari, A.; Marklund, E. G.; Collier, M. P.; Chandler, S. A.; Tkachenko, O.; Allen, J.; Crispin, M.; Billington, N.; Takagi, Y.; Sellers, J. R.; Eichmann, C.; Selenko, P.; Frey, L.; Riek, R.; Galpin, M. R.; Struwe, W. B.; Benesch, J. L. P.; Kukura, P. Quantitative Mass Imaging of Single Molecules. *Science (80-.).* **2018**, *360*, 423–427.

(51) Kao, A.; Chiu, C.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. Development of a Novel Cross-Linking Strategy for Fast and Accurate Identification of Cross-Linked Peptides of Protein Complexes. *Mol. Cell. Proteomics* **2011**, *10*, 1–17.

(52) Müller, M. Q.; Dreiocker, F.; Ihling, C. H.; Schäfer, M.; Sinz, A. Cleavable Cross-Linker for Protein Structure Analysis: Reliable Identification of Cross-Linking Products by

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Tandem MS. *Anal. Chem.* **2010**, *82*, 6958–6968.

(53) Iacobucci, C.; Götze, M.; Ihling, C. H.; Piotrowski, C.; Arlt, C.; Schäfer, M.; Hage, C.; Schmidt, R.; Sinz, A. A Cross-Linking/Mass Spectrometry Workflow Based on MS-Cleavable Cross-Linkers and the MeroX Software for Studying Protein Structures and Protein–Protein Interactions. *Nat. Protoc.* **2018**, *13*, 2864–2889.

(54) Shevchenko, A.; Tomas, H.; Havliš, J.; Olsen, J. V.; Mann, M. In-Gel Digestion for Mass Spectrometric Characterization of Proteins and Proteomes. *Nat. Protoc.* **2007**, *1*, 2856–2860.

(55) Götze, M.; Pettelkau, J.; Fritzsche, R.; Ihling, C. H.; Schäfer, M.; Sinz, A. Automated Assignment of MS/MS Cleavable Cross-Links in Protein 3d-Structure Analysis. *J. Am. Soc. Mass Spectrom.* **2014**, *26*, 83–97.

(56) Helten, A.; Koch, K. W. Calcium-Dependent Conformational Changes in Guanylate Cyclase-Activating Protein 2 Monitored by Cysteine Accessibility. *Biochem. Biophys. Res. Commun.* **2007**, *356*, 687–692.

(57) Fraczkiewicz, R.; Braun, W. Exact and Efficient Analytical Calculation of the Accessible Surface Areas and Their Gradients for Macromolecules. *J. Comput. Chem.* **1998**, *19*, 319–333.

(58) Ali, S.; Hassan, M.; Islam, A.; Ahmad, F. A Review of Methods Available to Estimate Solvent-Accessible Surface Areas of Soluble Proteins in the Folded and Unfolded States. *Curr. Protein Pept. Sci.* **2014**, *15*, 456–476.

(59) Hanić, M.; Frederiksen, A.; Schuhmann, F.; Solov'yov, I. A. On the Energetic Differences of Avian Cryptochromes 4 from Selected Species. *Eur. Phys. J. D* **2022**, *76*.

(60) Schuhmann, F.; Korol, V.; Solov'yov, I. A. Introducing Pep McConst—A User-Friendly Peptide Modeler for Biophysical Applications. *J. Comput. Chem.* **2021**, *42*, 572–580.

(61) Chen, R.; Weng, Z. A Novel Shape Complementarity Scoring Function for Protein-Protein Docking. *Proteins Struct. Funct. Genet.* **2003**, *51*, 397–408.

(62) Pierce, B. G.; Hourai, Y.; Weng, Z. Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library. *PLoS One* **2011**, *6*, 1–6.

(63) Pierce, B. G.; Wiehe, K.; Hwang, H.; Kim, B. H.; Vreven, T.; Weng, Z. ZDOCK Server: Interactive Docking Prediction of Protein-Protein Complexes and Symmetric Multimers. *Bioinformatics* **2014**, *30*, 1771–1773.

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

(65) Pierce, B.; Tong, W.; Weng, Z. M-ZDOCK: A Grid-Based Approach for C_n Symmetric Multimer Docking. *Bioinformatics* **2005**, *21*, 1472–1478.

(66) Zhang, Y. I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinformatics* **2008**, *9*, 1–8.

(67) Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A Unified Platform for Automated

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Protein Structure and Function Prediction. *Nat. Protoc.* **2010**, *5*, 725–738.

(68) Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein Structure and Function Prediction. *Nat. Methods* **2014**, *12*, 7–8.

(69) Yang, J.; Zhang, Y. I-TASSER Server: New Development for Protein Structure and Function Predictions. *Nucleic Acids Res.* **2015**, *43*, 174–181.

(70) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 1–33.

(71) Korol, V.; Husen, P.; Sjulstok, E.; Nielsen, C.; Friis, I.; Frederiksen, A.; Salo, A. B.; Solov'yov, I. A. Introducing VIKING: A Novel Online Platform for Multiscale Modeling. *ACS Omega* **2020**, *5*, 1254–1260.

(72) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. J. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Sidechain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(73) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulation. *J. Comput. Chem.* **2004**, *25*, 1400–1415.

(74) Lüdemann, G.; Solov'yov, I. A.; Kubař, T.; Elstner, M. Solvent Driving Force Ensures Fast Formation of a Persistent and Well-Separated Radical Pair in Plant Cryptochrome. *J. Am. Chem. Soc.* **2015**, *137*, 1147–1156.

(75) Solov'yov, I. A.; Domratcheva, T.; Schulten, K. Separation of Photo-Induced Radical Pair in Cryptochrome to a Functionally Critical Distance. *Sci. Rep.* **2014**, *4*, 1–8.

(76) Sjulstok, E.; Olsen, J. M. H.; Solov'yov, I. A. Quantifying Electron Transfer Reactions in Biological Systems: What Interactions Play the Major Role? *Sci. Rep.* **2015**, *5*, 1–11.

(77) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.

(78) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(79) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(80) Sousa, S. F.; Neves, R. P. P.; Waheed, S. O.; Fernandes, P. A.; Ramos, M. J. Structural and Mechanistic Aspects of S-S Bonds in the Thioredoxin-like Family of Proteins. *Biol. Chem.* **2019**, *400*, 575–587.

(81) Kabsch, W. A Discussion of the Solution for the Best Rotation to Relate Two Sets of

bioRxiv preprint doi: https://doi.org/10.1101/2023.04.11.536350; this version posted April 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Vectors. *Acta Crystallogr. Sect. A* **1978**, *A34*, 827–828.

(82) Prinz, W. A.; Åslund, F.; Holmgren, A.; Beckwith, J. The Role of the Thioredoxin and Glutaredoxin Pathways in Reducing Protein Disulfide Bonds in the *Escherichia Coli* Cytoplasm. *J. Biol. Chem.* **1997**, *272*, 15661–15667.

(83) Jakob, U.; Eser, M.; Bardwell, J. C. A. Redox Switch of Hsp33 Has a Novel Zinc-Binding Motif. *J. Biol. Chem.* **2000**, *275*, 38302–38310.

(84) Choi, H.; Kim, S.; Mukhopadhyay, P.; Cho, S.; Woo, J.; Storz, G.; Ryu, S. Structural Basis of the Redox Switch in the OxyR Transcription Factor. **2001**, *105*, 103–113.

(85) Wang, Q.; Lin, C. Mechanisms of Cryptochrome-Mediated Photoresponses in Plants. *Annu. Rev. Plant Biol.* **2020**, *71*, 103–129.

(86) Schmalen, I.; Reischl, S.; Wallach, T.; Klemz, R.; Grudziecki, A.; Prabu, J. R.; Benda, C. Interaction of Circadian Clock Proteins CRY1 and PER2 Is Modulated by Zinc Binding and Disulfide Bond Formation. *Cell* **2014**, *157*, 1203–1215.