ORIGINAL PAPERS

Key interactions with deazariboflavin cofactor for light-driven energy transfer in *Xenopus* (6–4) photolyase

Ayaka Morimoto[1] · Yuhei Hosokawa[1] · Hiromu Miyamoto[1] · Rajiv Kumar Verma[1,3] · Shigenori Iwai[1] · Ryuma Sato[2,4] · Junpei Yamamoto[1]

Abstract

Photolyases are flavoenzymes responsible for light-driven repair of carcinogenic crosslinks formed in DNA by UV exposure. They possess two non-covalently bound chromophores: flavin adenine dinucleotide (FAD) as a catalytic center and an auxiliary antenna chromophore that harvests photons and transfers solar energy to the catalytic center. Although the energy transfer reaction has been characterized by time-resolved spectroscopy, it is strikingly important to understand how well natural biological systems organize the chromophores for the efficient energy transfer. Here, we comprehensively characterized the binding of 8-hydroxy-7,8-didemethyl-5-deazariboflavin (8-HDF) to *Xenopus* (6–4) photolyase. In silico simulations indicated that a hydrophobic amino acid residue located at the entrance of the binding site dominates translocation of a loop upon binding of 8-HDF, and a mutation of this residue caused dysfunction of the efficient energy transfer in the DNA repair reaction. Mutational analyses of the protein combined with modification of the chromophore suggested that Coulombic interactions between positively charged residues in the protein and the phenoxide moiety in 8-HDF play a key role in accommodation of 8-HDF in the proper direction. This study provides a clear evidence that *Xenopus* (6–4) photolyase can utilize 8-HDF as the light-harvesting chromophore. The obtained new insights into binding of the natural antenna molecule will be helpful for the development of artificial light-harvesting chromophores and future characterization of the energy transfer in (6–4) photolyase by spectroscopic studies.

Keywords Drug binding · Photolyase · Light-driven energy transfer · Molecular dynamics · DNA repair

1 Introduction

Natural biological systems utilize light as an external energy source to drive chemical reactions in organisms, e.g., water oxidation in photosystem II (PSII) [1], signal transduction by rhodopsin [2] and photoactive yellow protein [3], and light-driven DNA repair of the UV-induced DNA damage by photolyases (PLs) [4, 5]. They commonly possess light-absorbing chromophores, and the excited state of the chromophore induces light-induced electron transfer or *cis–trans* isomerization of the chromophore, triggering their biological functions. On the other hand, the functions of these proteins strongly rely on the absorption window of the chromophore. To expand the limitation, some biological systems utilize natural auxiliary chromophores that either harvest the light energy with better efficiency than the catalytic chromophore or fill the absorption window gap of the catalytic chromophore [6]. In both cases, the light energy absorbed by the auxiliary chromophore is transferred to the

✉ Junpei Yamamoto
yamamoto@chem.es.osaka-u.ac.jp

1   Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

2   Center for Biosystems Dynamics Research, RIKEN, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

3   Present Address: Synthetic Organic Chemistry Laboratory, RIKEN Cluster for Pioneering Research, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

4   Present Address: Cellular and Molecular Biotechnology Research and Institute, National Institute of Advanced Industrial Science and Technology (AIST), 2-3-26 Aomi, Koto-ku, Tokyo 135-0064, Japan

functional chromophore via the Förster resonance energy transfer (FRET) mechanism, achieving the more efficient light-induced biological functions. Inspired by the natural systems, enhancement of the light-absorbing abilities by exogenous artificial chromophores has been investigated and attracted much interests [7–9]. A large part of the studies has been devoted to trace quenching of the fluorescence of the donor and/or rise of absorption of the excited state of the acceptor by time-resolved absorption/fluorescence spectroscopy [10–14]. Considering the Förster theory, however, the FRET efficiency between two chromophores highly depends on the amplitude, distance, and orientation of the transition dipole moment of the respective chromophore, in addition to the spectral overlap between the absorption of the acceptor and the fluorescence of the donor [15]. Toward the development of artificial molecules that enhance the light-driven reactions via FRET, it is strikingly important to understand how the enzymes utilize the auxiliary chromophore in the natural system.

PLs, which exist in all domains of life and are responsible for light-driven repair of carcinogenic DNA damage such as cyclobutene pyrimidine dimers (CPDs) or pyrimidine(6–4) pyrimidone photoproducts (6–4PPs), are good candidates to understand the light-utilizing mechanism with the auxiliary chromophore. They are soluble proteins harboring a sole light-absorbing catalytic center, flavin adenine dinucleotide (FAD), and one auxiliary chromophore enhancing the DNA repair activity via FRET (Fig. 1). The key reaction in the DNA repair by PLs that specifically repair either CPD or 6–4PP (CPD-PL or (6–4)PL, respectively) is the electron transfer from the excited state of catalytically active fully-reduced form of FAD (FADH$^-$) to the photoproduct in DNA, which induces the bond cleavage and the rearrangement to the intact pyrimidine nucleobases [16].

Contrary to the exclusively conserved catalytic chromophore, several types of the auxiliary chromophore are known to function as the light-harvesting chromophore (LHC) in PLs. The examples are 5,10-methenyltetrahydrofolate (MTHF) in bacterial PL [17], FAD or flavin mononucleotide (FMN) in thermophilic bacterial PLs [20, 21], 6,7-dimethyl-8-ribityllumazine (DMRL) in bacterial (6–4)PLs [19], and 8-hydroxy-7,8-didemethyl-5-deazariboflavin (8-HDF) in class II CPD-PL [22], class I PL in cyanobacteria [23], algae cryptochrome/photolyase families (CPF) [18, 24], and insect (6–4)PL [24]. Regardless of the chemical structures of LHCs, their fluorescence overlaps the FADH$^-$ absorption, leading to the enhanced light-driven DNA repair via FRET. Among the known LHCs for PLs, 8-HDF is the only molecule that cannot be synthesized in *Escherichia coli* due to the lack of the endogenous 8-HDF biosynthetic pathway [22], and therefore, production of the 8-HDF-binding PLs in *E. coli* typically yields the apo PLs without 8-HDF. This means that reconstruction of the PLs bearing an artificial
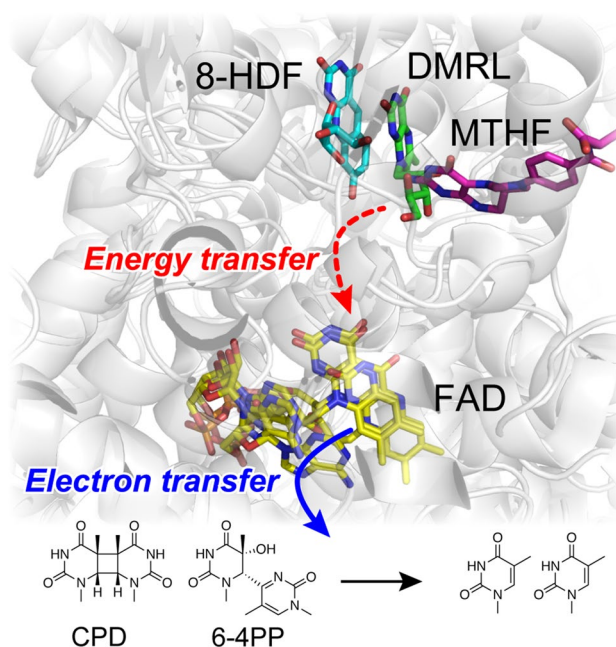

Fig. 1 Light-harvesting chromophores in the light-driven DNA repair by photolyases. The crystal structures of *Escherichia coli* CPD-PL (1DNP) [17], *Chlamydomonas reinhardtii* animal-like cryptochrome (6FN2) [18], and *Agrobacterium tumefaciens* (6–4)PL (4DJA) [19] were superimposed, and their auxiliary chromophores (MTHF, 8-HDF, and DMRL for 1DNP, 6FN2, and 4DJA, respectively) are shown

LHC is possible only for 8-HDF-binding PLs. Indeed, soaking the crystals of apoproteins of *Methanosarcina mazei* class II CPD-PL into the 8-HDF solution reportedly gave the crystals of 8-HDF bound holoenzymes [22]. Yet, the interactions dominating the 8-HDF uptake into the apo PLs remain obscure.

Here, we investigated the 8-HDF binding to *Xenopus laevis* (6–4)PL (*Xl*64), the LHC of which has not been identified so far. Molecular dynamics (MD) simulations and binding energy calculations in combination with experiments using *Xl*64 mutants and synthetic deazariboflavin molecules identified the key interactions, which are crucially important for the 8-HDF uptake and its light-harvesting function. This work thus provides a guideline for the molecular design of the artificial LHC in PLs.

2 Experimental methods

2.1 General

Primers for the PCR reaction were purchased from Thermo Fisher Scientific. Reagents for the synthesis of the 8-HDF derivatives were purchased from Tokyo Chemical Industry or Fuji-Film Wako Pure Chemical Industries. TLC analyses

were performed on Merck Silica gel 60 F254 plates, and the results were visualized by UV illumination at 254 nm. For column chromatography, Wakogel C-200 was used. HPLC analyses were performed either on a Shimadzu gradient-type analytical system equipped with a Shimadzu SPD-M10AVP photodiode-array detector or a Gilson gradient-type analytical system equipped with a Waters 2998 photodiode-array detector. A Waters μBondasphere C18 5 μm 300 Å column (3.9 × 150 mm) or a Waters XBridge C18 5 μm column (4.6 × 150 mm) were used on these systems, at a flow rate of 1.0 mL min^{-1}, with a linear gradient of acetonitrile in 0.1 M triethylammonium acetate (TEAA) (pH 7.0) generated over 20 min. Absorption and fluorescent spectra were measured on a PerkinElmer Lambda 35 UV/Vis spectrometer and a JASCO FP-6500 spectrofluorometer, respectively. The pH values were measured with a HORIBA pH meter F-71 equipped with a 9618S-10D electrode, after 3-point calibration.

2.2 Synthesis of 8-HDF and its 8-methoxy derivative

8-HDF was synthesized according to the previous report [25]. The aiming product was finally purified with HPLC using the XBridge column, without the TEAA buffer. The pooled solutions were combined and lyophilized. The yellowish powder was weighed with a microbalance, and dissolved in DMSO at a concentration of 20 mM. The 8-methoxy derivative was synthesized according to the procedure for 8-HDF, except that *m*-anisidine was used as a starting material, instead of *m*-aminophenol (see Supplementary Information for analytical data). The DMSO solutions of the compounds were used for the following experiments.

2.3 Site-directed mutagenesis

The genes of the F41A and R51A mutants of *Xl*64 were generated with the PrimeSTAR Mutagenesis Basal kit (Takara, Japan) using the kanamycin-resistant pET28a plasmid as a template, in which the *Xl*64 gene was integrated at the *Nde*I/*Xho*I site (pET28a-Xl64). The PCR primers (Primers 1–4) are summarized in Table S1. The obtained plasmids (pET28a-Xl64·F41A and pET28a-Xl64·R51A) were sequenced and used for the protein production.

For introduction of the K256A mutation, the plasmid reconstruction using the In-fusion HD cloning kit (Takara, Japan) was employed. The 5' upstream and 3' downstream of K256 of either pET28a-Xl64 or pET28a-Xl64 R51A were amplified with PrimeSTAR MAX DNA polymerase (Takara) using two sets of PCR primers (Primer 5–8 in Table S1). Noteworthy, Primers 6 and 7 contain the DNA sequences for the K256A mutation. The amplicons were purified with agarose gel. In addition, the pET28a-Xl64 plasmid was treated with *Nde*I and *Xho*I at 37 °C overnight, and the 5 kb

product was purified with agarose gel. The linearized vector and the 5' and 3' amplicons were assembled with the In-fusion HD cloning kit (Takara). The obtained plasmids containing the K256A mutation (pET28a-Xl64·K256A and pET28a-Xl64·R51A/K256A) were sequenced and used for the protein production.

2.4 Protein purification

Wild-type *Xl*64 and *Arabidopsis thaliana* (6–4)PL (*At*64) were prepared according to the previous reports using the ampicillin-resistant modified pET28a plasmid encoding either the *Xl*64 (pET28a'-Xl64) [26] or *At*64 gene [27], whereas the mutants of *Xl*64 (F41A, R51A, K256A, and R51A/K256A) were prepared using the kanamycin-resistant pET28a plasmids, which were prepared as mentioned above. Regarding the preparation of the *Xl*64 mutants, *E. coli* BL21 (DE3) star cells (Merck) were transformed with the pET28a plasmid bearing the mutant gene. A single colony was inoculated into 40 mL of a Luria-Broth (LB) medium with kanamycin (20 μg mL^{-1}), and the overnight culture grown at 37 °C was further inoculated into 2 L of a Terrific Broth (TB) medium with the antibiotics. The culture was shaken at 37 °C until OD$_{600}$ reached 0.8, and isopropyl β-D-1-thiogalactopyranoside (IPTG) was added at a final concentration of 20 μM. The culture was then shaken at 16 °C for 36 h. After harvest, the cells were suspended in a lysis buffer (50 mM NaH$_2$PO$_4$, 200 mM NaCl, and 5 mM imidazole, pH 8), and lysed with sonication. The mutant proteins were purified from the supernatant after the centrifugation of the cell lysate, via the TALON (Clontech) and Heparin (GE Healthcare) columns, as described previously. The purified enzymes were analyzed by SDS-PAGE and UV absorption spectroscopy, and their concentrations were measured by the Bradford analysis and the FAD absorbance at 450 nm using a molecular extinction coefficient of 11,200 M^{-1} cm^{-1}.

Reconstitution of *Xl*64 bound to biosynthetic 8-HDF was achieved using a pCDFDuet-1 plasmid, in which *Streptomyces coelicolor* FO synthase gene was integrated at the *Nde*I/*Xho*I site (pCDFDuet-*fbi*C) [22]. *E. coli* C41(DE3) pLysS competent cell (Lucigen) was transformed with pCDFDuet-*fbi*C and pET28a'-Xl64, and the transformant was selected on a LB plate containing ampicillin (50 μg mL^{-1}), chloramphenicol (37 μg mL^{-1}), and spectinomycin (50 μg mL^{-1}). A single colony was inoculated into 40 mL of the LB medium containing ampicillin (50 μg mL^{-1}) and spectinomycin (50 μg mL^{-1}), and the overnight culture grown at 37 °C was further inoculated into 2 L of the LB medium containing the antibiotics. The culture was shaken at 25 °C until OD$_{600}$ reached 0.8. After addition of IPTG at a final concentration of 1 mM, the culture was further shaken at 25 °C for 24 h. After harvest, the cell pellet was suspended in the lysis buffer, and were lysed by sonication. The protein

was purified from the supernatant after the centrifugation, in the same way as the mutants. The purified enzyme was analyzed by SDS-PAGE and UV/Vis absorption spectroscopy, and the concentration was evaluated by the Bradford assay. The co-expression experiment of *At*64 and FO synthase was also performed in the same way using the plasmids.

2.5 In vitro *reconstitution of chromophore-bound enzymes*

To reconstitute the chromophore-bound enzymes in vitro, 10 μM enzymes and 100 μM chromophores were mixed in 50 μL of buffer A (10 mM sodium phosphate, 100 mM NaCl, and 5% glycerol, pH 7), and the mixtures were incubated on ice for 1 h. The samples were loaded on Micro Bio-Spin 6 gel filtration columns (Bio-Rad), pre-equilibrated with buffer A, to remove the unbound chromophores. The eluate was used for the following experiments.

For qualitative binding analyses, the eluates of the in vitro reconstituted chromophore-bound enzymes were heated at 80 °C for 5 min to isolate the bound FAD and antenna chromophores from the proteins. The mixture was filtrated through the Ultrafree centrifugal filter unit with 0.45 μm pore size (Millipore), and the filtrate was analyzed by HPLC, using a linear gradient of 7–13% acetonitrile in 0.1 M TEAA. The same procedure was applied to the in vivo reconstituted chromophore-bound enzymes. The peak area of FAD detected at 450 nm and that of 8-HDF/8-MDF detected at 420/395 nm were divided by molar extinction coefficients (11,200 for FAD, 23,000 for 8-HDF, and 13,300 for 8-MDF). From these values, the molar ratio of FAD/8-HDF and FAD/8-MDF were obtained. The experiments were performed in triplicate.

For dissociation analyses, the in vitro reconstituted 8-HDF bound *Xl*64 proteins were diluted with buffer A, and the samples were placed in a cuvette chilled at 10 °C. The absorption spectra of the samples were measured at an interval of 30 min, and the change of the absorption maximum was evaluated.

2.6 Spectroscopic analyses

UV/Vis absorption and fluorescence spectra of 8-HDF and 8-MDF (10 μM) were measured in 10 mM sodium phosphate buffer at various pH. For fluorescence measurements, the excitation wavelengths of 8-HDF and 8-MDF were set to 420 and 395 nm, respectively. Titration of the proteins into the 8-HDF solution was performed as follows; 15 μM of 8-HDF or 8-MDF dissolved in a buffer (200 μL, 10 mM sodium phosphate (pH 7.0), 100 mM NaCl, and 5% glycerol) was placed in a cuvette, and an aliquot (2 μL) of the concentrated protein solution (200 μM) were added into the cuvette. After 10 min incubation on ice, a spectrum was

measured. The titration and measurement processes were repeated 10 times.

Fluorescence spectra of the 8-HDF bound enzymes, reconstituted as described above, were measured in the buffer (200 μL, 500 nM enzymes, 10 mM sodium phosphate (pH 7.0), 100 mM NaCl, and 5% glycerol), in a range of 435–600 nm.

2.7 DNA repair experiment

Repair of 6–4PP by the *Xl*64 proteins was performed according to the previous report [9]. Briefly, the purified enzyme (2 μM) in reaction buffer (94 μL, 20 mM sodium phosphate, 500 mM NaCl, and 10% glycerol, pH 7.5) was placed in a 10×2×8 mm (length×width×height) inner volume quartz cuvette (Starna, 16.160-F/4/Q/10 GL 14/2/Z15) and purged with nitrogen through PTFE-coated silicone and rubber septa. L-Cysteine was added to the solution in a glovebox to give a final concentration of 5 mM. After 10 min irradiation with UV-light at 365±5 nm from a UV-LED light source LC-LIV3 (Hamamatsu Photonics) through a 10×8 mm window, to photoreduce FAD to fully-reduced flavin adenine dinucleotide (FADH⁻), the single-stranded 8-mer substrate, d(CGAT(6–4)TGAT), where T(6–4)T represents 6–4PP, was added to the solution at a final concentration of 40 μM under dark and anaerobic conditions. The final volume of the sample solution was 100 μL. The mixture was illuminated on ice with white light (430–800 nm) from a MAX-150 xenon lamp (Asahi Spectra) through the 10×8 mm window for a total illumination time of 24 s, with an interval resting time of 3 min. After each illumination, aliquots (5 μL) of the mixture were sampled through the septa and deproteinized with phenol/chloroform/isoamyl alcohol (25:24:1). After recovery of DNA by ethanol precipitation, the samples were analyzed by HPLC with an acetonitrile gradient from 5 to 13% to determine the numbers of repaired substrates per enzyme.

2.8 Setups for theoretical calculations

As the structure of *Xl*64 has not been reported so far, we used *Drosophila melanogaster* (6–4)PL (*Dm*64) structures with or without the 8-HDF cofactor (3CVV [24] and 3CVU [28], referred to as holo*Dm*64 and apo*Dm*64, respectively). Note that *Xl*64 and *Dm*64 share 58% identity and 70% homology, and F41, R51, and K256 in *Xl*64 correspond to I50, R60, and K266 in *Dm*64, respectively. We also created the 8-HDF-bound 3CVU structure (apo*Dm*64•HDF) using a docking simulation program in the Schrödinger suite. Also, structures of apo*Dm*64-I50A•HDF and apo*Dm*64-R60A•HDF were created by editing the PDB file of apo*Dm*64•HDF.

To model the calculation systems, tLeap module in AMBER16 program package [29] was used to each structure. In this study, the AMBER14SB force field [30] for the

standard amino acid residues and the TIP3P water model [31] for water molecules were adopted, while self-created force fields were applied to FAD and 8-HDF. We created the force fields of FAD and 8-HDF as follows; first, hydrogen atoms were added to each chromophore in the 3CVV file using a Winmostar software, and then the geometry optimization with B3LYP/6-31G(d) level of theory and the electronic structure calculation with MP2/6–311++G(2d,2p) level of theory were performed using Gaussian16 [32]. The Gaussian output data were converted to AMBER force field using the antechamber module in the AMBER16 program package and a general amber force field [33]. Note that the restrained electrostatic potential charge was applied. Each initial system was solvated in a rectangular box containing TIP3P water molecules within 12 Å from the protein, and counter ions (Na^+) were added to neutralize the system. Finally, the AMBER topology file and the AMBER coordinate file were converted to the GROMACS topology file and the GROMACS coordinate file, using ACPYPE [34].

The calculation system for *At*64 was also created using the reported coordinates (3FY4) [35], in the same way as for *Dm*64.

2.9 Molecular dynamics simulations

All MD simulations were performed using the GROMACS version 5.1.4 program [36]. First, the energy minimization of each system was performed with 5000 steps, followed by the MD simulation for 10 ps in the NVT ensemble ($T = 5$ K). The temperature was subsequently increased from 5 to 298 K for 90 ps in the NPT ensemble ($P = 1$ bar and $T = 5, 50, 100, 150, 200, 225, 250, 275,$ and 298 K). Then, 500 ns MD simulation in the NPT ensemble ($P = 1$ bar and $T = 298$ K) was performed for holo*Dm*64, apo*Dm*64, apo*Dm*64●HDF, apo*Dm*64-I50A●HDF, and *At*64, while 600 ns for apo*Dm*64-R60A●HDF. The temperature and pressure were kept constant using the Nose–Hoover thermostat [37] and Berendsen barostat [38], respectively. The linear constraint solver (LINCS) algorithm [39] was applied to all covalent bonds, with an integration time step of 2.0 fs. The long-range Coulomb interactions were treated with the particle mesh Ewald method [40] and the direct space cutoff distance was set to 10 Å. The van der Waals interactions were calculated using a switched cutoff between 8 and 10 Å.

2.10 Binding free energy estimation with molecular mechanics/Poisson–Boltzmann surface area (MM-PBSA)

To estimate the binding free energy and the contribution of each amino acid residues in the binding pocket, we performed molecular mechanics/Poisson–Boltzmann surface area (MM-PBSA) calculations (see Supplementary theory) using the g_mmpbsa tool [41]. In the MM-PBSA calculation process using g_mmpbsa, the polar solvation energy was calculated by an assisted Poisson–Boltzmann solver (APBS) program [42], while the apolar solvation energy was calculated by Solvent-accessible surface area (SASA) model [43].

3 Results and discussions

3.1 Molecular dynamics simulations

Among the (6–4)PLs, *Dm*64 [24, 28] and *Chlamydomonas reinhardtii* animal-like cryptochrome (*Cr*aCRY) [18] are the only enzymes known to bind 8-HDF. To gain insights into the capabilities of *Xl*64 to bind 8-HDF, the reported *Dm*64 structures with or without 8-HDF (apo*Dm*64 [28] or holo*Dm*64 [24], respectively) were used for the initial in silico screening of the interactions in the 8-HDF binding. It is noteworthy that *Xl*64 and *Dm*64 share high homology in the amino acid sequence (58% identity and 70% similarity), and the amino acid residues located within the 5 Å distance from 8-HDF in holo*Dm*64 (Fig. 2A, B) are well conserved in *Xl*64, except the 3 amino acids in the loop moiety. Interestingly, in the crystal structure of holo*Dm*64, this loop in apo*Dm*64 was rearranged into an α-helix and translocated onto the 8-HDF π–plane (Fig. 2C), as if a lid in the binding site is closed upon the 8-HDF uptake. Especially, the Ile50 side chain in holo*Dm*64 is located over 8-HDF at a 3 Å distance (Fig. 2B), indicating a role in the loop translocation. To reproduce the dynamic transition of the loop from the open state to the closed state, we performed molecular dynamics simulations. First, a docking simulation of 8-HDF with apo*Dm*64 was performed with a Schrödinger docking simulator. The output structure (apo*Dm*64●HDF) exhibited the successful 8-HDF incorporation into a position similar to the binding site in holo*Dm*64 (Fig. 2C), while the loop remained in the open state. Using the apo*Dm*64●HDF coordinates as an initial structure, in addition to the holo*Dm*64 coordinates, 500 ns MD simulations were performed. Interestingly, the loop in question came close to the HDF moiety within the 5 ns simulation time, and stayed there during the remaining simulation time (Fig. 2D). Although the loop moiety in apo*Dm*64●HDF was not completely transformed to the α-helix as observed in the holo*Dm*64 crystal structure, these observations indicate that the transition is caused by the 8-HDF uptake in the binding site.

To estimate the roles of the amino acid residues in the 8-HDF binding pocket, molecular mechanics/Poisson–Boltzmann surface area (MM-PBSA) and pocket volume measurer (POVME) calculations using the MD trajectories of holo*Dm*64 and apo*Dm*64●HDF were performed. MM-PBSA analysis (detailed data of which are shown in Fig. S2) indicated that 6 amino acid side chains (Arg13, Lys14, Phe44,
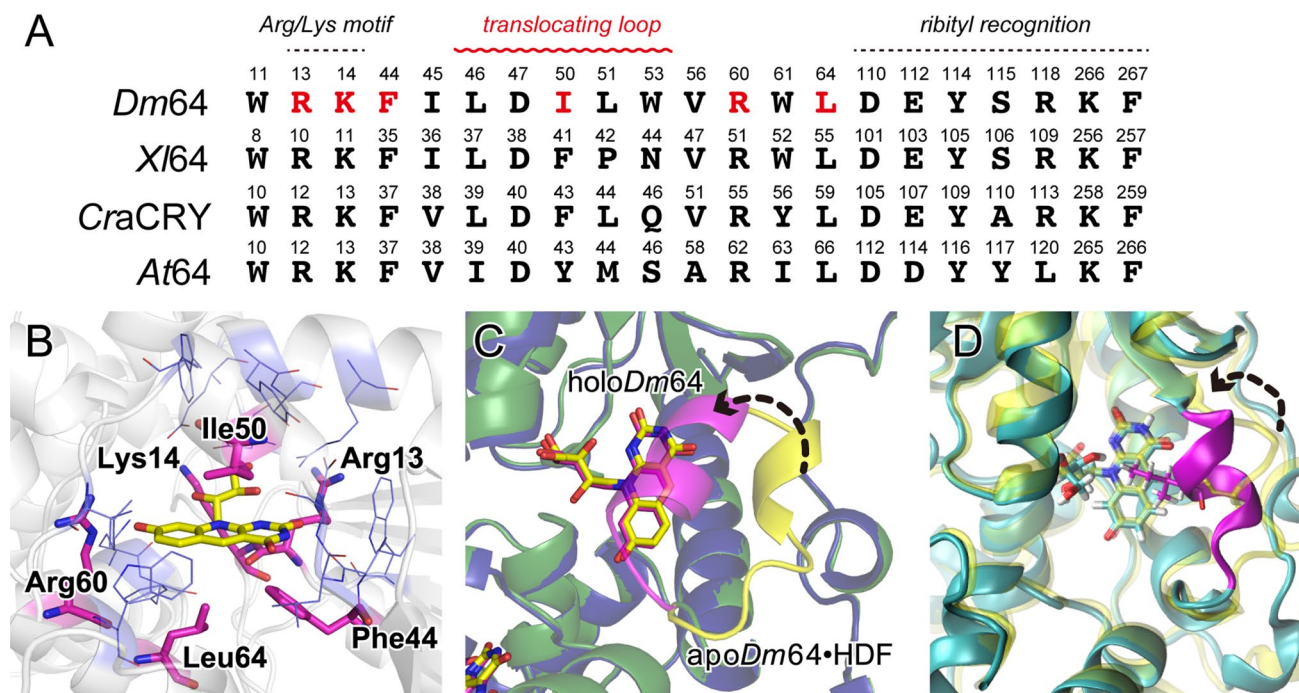
A

		Arg/Lys motif				translocating loop								ribityl recognition						
Dm64	11	13	14	44	45	46	47	50	51	53	56	60	61	64	110	112	114	115	118	266	267
	W	R	K	F	I	L	D	I	L	W	V	R	W	L	D	E	Y	S	R	K	F
Xl64	8	10	11	35	36	37	38	41	42	44	47	51	52	55	101	103	105	106	109	256	257
	W	R	K	F	I	L	D	F	P	N	V	R	W	L	D	E	Y	S	R	K	F
CraCRY	10	12	13	37	38	39	40	43	44	46	51	55	56	59	105	107	109	110	113	258	259
	W	R	K	F	V	L	D	F	L	Q	V	R	Y	L	D	E	Y	A	R	K	F
At64	10	12	13	37	38	39	40	43	44	46	58	62	63	66	112	114	116	117	120	265	266
	W	R	K	F	V	I	D	Y	M	S	A	R	I	L	D	D	Y	Y	L	K	F

B (Ile50, Lys14, Arg13, Arg60, Leu64, Phe44)

C holoDm64, apoDm64•HDF

D

Fig. 2 The 8-HDF binding site of (6–4)PLs. **a** Amino acid residues located within 5 Å distance from 8-HDF in the holoDm64 structures and the corresponding amino acid residues in Xl64, Chlamydomonas reinhardtii animal-like cryptochrome (CraCRY), and At64. Their multiple sequence alignment in this region is shown in Fig. S1. The residues in red exhibited free binding energies to 8-HDF smaller than –1.5 kcal/mol. **b** A structural view of the 8-HDF binding site in the crystal structure of holoDm64. The amino acid residues shown in Panel A are visualized in thin or thick lines, the latter of which represent the residues with < –1.5 kcal/mol binding energies. **c** An overlay of the input structures of holoDm64 (blue) and apoDm64•HDF (green), in which the translocating loops are shown in magenta and yellow, respectively. **d** An overlay of the input structure of apoDm64•HDF (yellow) and a snapshot of its MD trajectory at 23 ns (light green), in which the translocating loop from Leu46 to Trp53 are highlighted in magenta

Ile50, Arg60, and Leu64 in Fig. 2B) exhibited attractive total binding free energies to 8-HDF smaller than − 1.5 kcal/mol, among the amino acids located within the 5 Å distance from 8-HDF in the holoDm64 structure (Fig. 2A). Phe44 and Leu64 participated in the formation of the hydrophobic binding pocket whereas the Arg13 and Lys14 were involved in the recognition of hydrophilic ribityl chain of 8-HDF, associated with their surrounding amino acids. Arg60 is located in the close vicinity to the hydroxy group of 8-HDF, suggesting formation of the electrostatic and/or hydrogen bonding interaction. The binding free energy for Ile50 in holoDm64 was − 3.3 kcal/mol, indicating a CH–π interaction between Ile50 and 8-HDF. After the MD simulation of apoDm64•HDF, the alkyl side chain of Ile50 is also located over 8-HDF with the binding free energy of − 3.0 kcal/mol (Fig. 2D), indicating that the attractive interaction due to the CH–π interaction likely dominates the translocation of the loop, to form the closed state.

In the POVME calculations, the 8-HDF molecule in the trajectories was removed, and the pocket volumes (POVs) within 7 Å from the center of mass of 8-HDF were calculated for each output frame. Figure 3A shows the distribution of the POV during the 500 ns MD simulation. The POVs of apoDm64 and holoDm64 were found to be 502 and 193 Å³, respectively. The shrink of the POV upon the 8-HDF binding was caused by the translocation of the loop. The POV of apoDm64•HDF, where the 8-HDF binding site was in the open state, was reduced by half (247 Å³) as compared to the open state in apoDm64, indicating that the loop translocation was caused by the capture of the 8-HDF. To identify the role of Ile50, the Ile50 residue in the initial structure of apoDm64•HDF was mutated into alanine (apoDm64-I50A•HDF), and 500 ns MD simulation was performed (Run1). Interestingly, the POV of apoDm64-I50A•HDF exhibited a bimodal distribution at 310 and 438 Å³ (Fig. 3A), which correspond to the closed and open states, respectively, as judged from the snapshot of the trajectories (Fig. 3C, D). The change of the POVs during the 500 ns simulation time (Fig. 3B) showed that the POV of apoDm64-I50A•HDF was unstably fluctuated in the first 300 ns simulation, and then fell into the closed state in the last 200 ns. To reproduce the simulation, an additional MD simulation of apoDm64-I50A•HDF was performed using the same initial structure (Run2). The bimodal POV distribution of the open and closed states was reproduced for Run2 (Fig. 3A), with the inverse population of each state as compared to that for
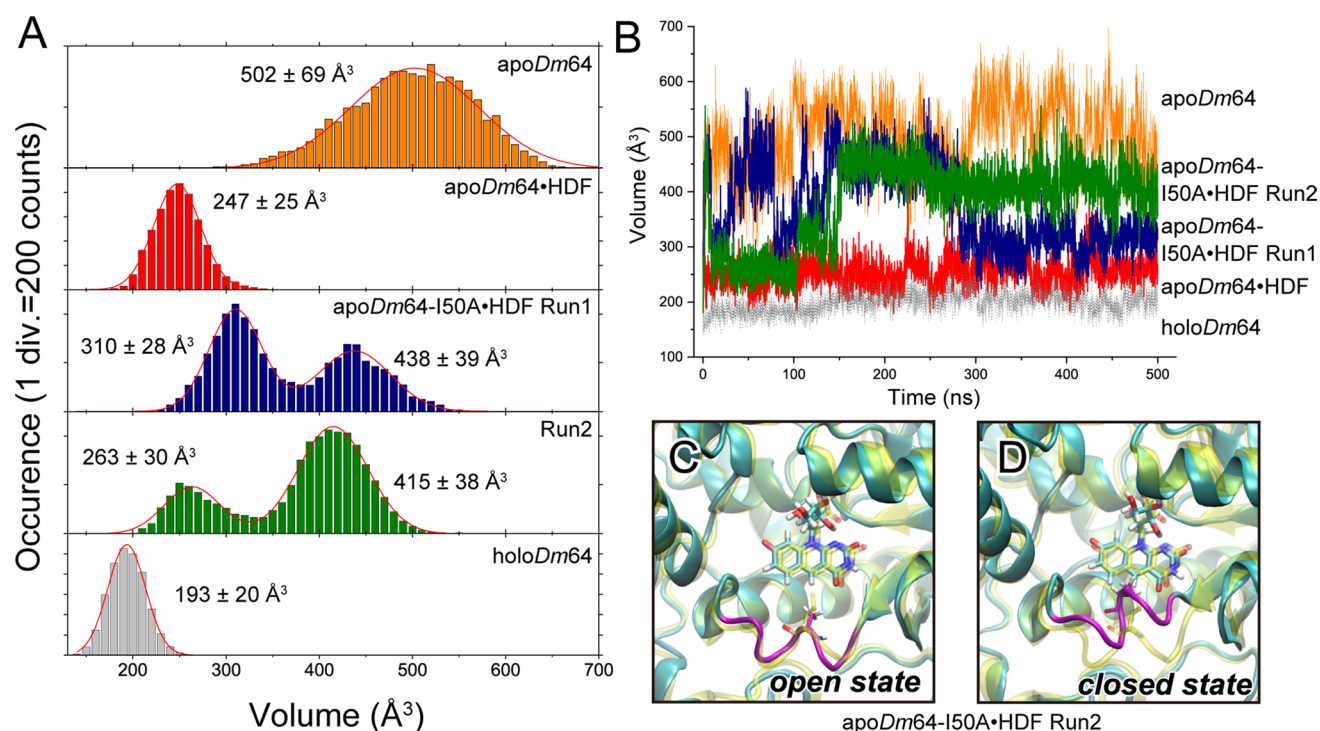
Fig. 3 The change of the 8-HDF binding site during the 500 ns MD simulations. **a** Distribution of the POVs in the MD simulations for apo*Dm*64 (orange), apo*Dm*64•HDF (red), apo*Dm*64-I50A•HDF (blue and green for Run1 and Run2, respectively), and holo*Dm*64 (gray). In the POVME analysis, the POV for each snapshot was estimated, and the time development of the POV was obtained, which is shown in Panel B. The number of the frames exhibiting the POV value within the same 10 Å³ range was counted among the 5000 snapshots, and the histograms were created. The histograms were then fitted with either a gauss function or a bimodal gauss function, and the results are shown in red line. The center of the distribution ± σ values obtained by the fitting are also shown. **b** Time development of the POV during the 500 ns MD simulations. **c, d** Snapshots of MD trajectories of apo*Dm*64-I50A•HDF Run2 at 5 ns (**c**) and 23 ns (**d**). To visualize the similarity and difference between the input structure and the trajectories, the trajectories (light green) were superimposed with the input structure of apo*Dm*64•HDF (yellow). The translocating loop in question in the snapshots is highlighted in magenta, in the same manner as in Fig. 2D

Run1. The time development of the POV for Run2 indicated that the binding site was in the open state during the last 350 ns (Fig. 3B). Although two runs gave a different trend in terms of the POV development, these observations suggest that the presence of Ile50 ensures the loop translocation and locking the lid of the 8-HDF binding site. Indeed, an MD simulation of R60A mutant (apo*Dm*64-R60A•HDF), which possesses the Ile residue in question, exhibited a POV development similar to that of apo*Dm*64•HDF, except much clearer POV shrinking at the initial stage of the simulation for this mutant than for wild type (Fig. S3A).

3.2 Binding of 8-HDF to *Xenopus* (6–4) photolyase

As shown in Fig. 2A, the presumed 8-HDF binding site of *Xl*64 is well conserved, except the 3 amino acid residues in the translocating loop, suggesting that 8-HDF would also bind to *Xl*64. To test the possibility, 8-HDF was chemically synthesized according to the previous literature [25], and in vitro reconstitution of the (6–4)PL•HDF complex was investigated in a qualitative manner, using gel filtration columns. In addition, to gain insights into the role of the functional group in 8-HDF, an 8-HDF derivative in which the hydroxy group was replaced with a methoxy group, namely 8-methoxy-7,8-didemethyl-5-deazariboflavin (8-MDF), was also synthesized (Fig. 4A, see Supplementary Information), and its binding ability to the (6–4) PLs were evaluated. After incubation of the recombinant *Xl*64 and plant *Arabidopsis thaliana* (6–4)PL (*At*64) with either 8-HDF or 8-MDF for 1 h, followed by passing through the gel filtration column, the eluted proteins were denatured by heating to isolate the bound chromophores. The supernatants were analyzed by HPLC (Fig. S4), and the molar ratios of 8-HDF/8-MDF and FAD are shown in Fig. 4B. Obviously, synthetic 8-HDF was detected in addition to FAD in the *Xl*64 sample, while 8-HDF was fully trapped with the gel filtration column in the absence of *Xl*64 (Fig. S4), indicating that 8-HDF was bound to *Xl*64. On the other hand, the amount of 8-MDF bound to *Xl*64 was significantly small. Since the hydroxy group is lacking
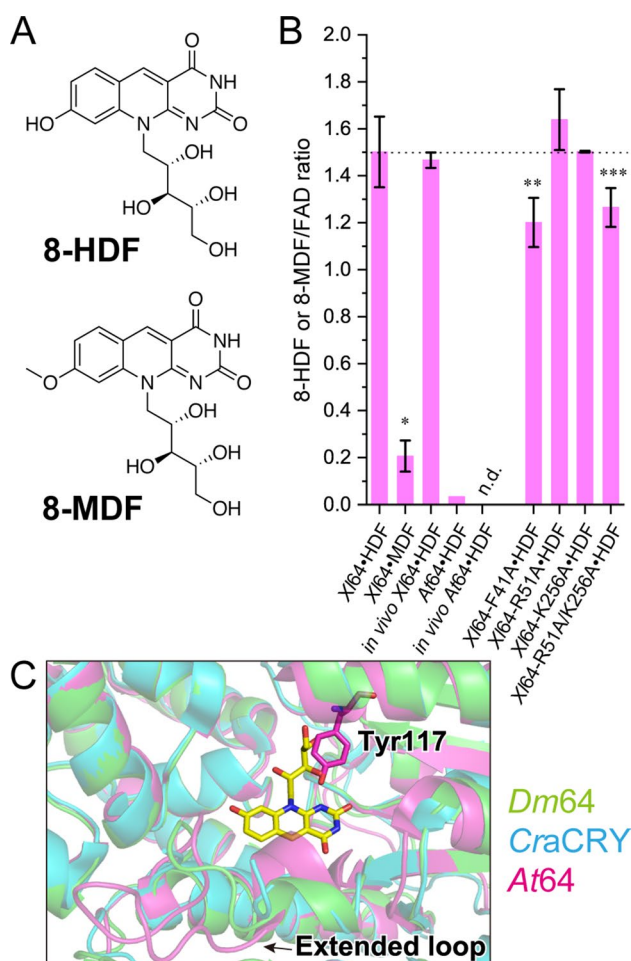
Fig. 4 8-HDF uptake in *Xenopus* and *Arabidopsis* (6–4)PLs. **a** Chemical structures of 8-HDF and 8-MDF. **b** 8-HDF or 8-MDF uptake in *Xl*64 and *At*64. After incubation of (6–4) PLs in the presence of 8-HDF or 8-MDF, followed by passing through a gel filtration column, the proteins were denatured by heating, and the supernatants were analyzed by HPLC (Fig. S4). Regarding the (6–4) PLs prepared by the co-expression system using the pCDFDuet-*fbi*C plasmid, the purified proteins were first passed through the gel filtration column. The recovered proteins were denatured by heating, and the supernatants were analyzed by HPLC. The experiments using the *Xl*64 proteins were performed in triplicate ($n=3$), and the mean values and the standard deviations of the 8-HDF or 8-MDF/FAD molar ratios are shown. The statistical significance compared to in vitro reconstituted *Xl*64•HDF complex was set to $P<0.05$. The asterisks indicate $P<0.001$ (*), $P<0.05$ (**), and $P<0.1$ (***). Regarding *At*64, the experiment was performed once as the synthetic and biosynthetic 8-HDF were hardly observable (n.d. represents not detected, see trace (F) in Fig. S4). **c** A superimposed view of the 8-HDF-binding site in the reported crystal structures of *Dm*64 (green, PDB 3CVU) [28], *Cr*aCRY (cyan, PDB 6FN2) [18], and *At*64 (magenta, PDB 3FY4) [43]. 8-HDF in the *Cr*aCRY structure is highlighted in yellow. Obviously, Tyr117 in *At*64 overlaps with the ribityl moiety of 8-HDF

in 8-MDF, these results imply a role of this functional group in the 8-HDF binding.

Apart from the results of *Xl*64, the presence of 8-HDF could not be observed for the *At*64 sample. To corroborate

the results of *Xl*64 and *At*64, in vivo reconstitution of the (6–4)PL•HDF complex of *Xl*64 (holo*Xl*64) was performed, using a plasmid coding a gene of FO synthase, an enzyme responsible for the biosynthesis of 8-HDF in *Streptomyces coelicolor* [22]. As observed in the in vitro reconstitution experiment, biosynthetic 8-HDF was detected in addition to FAD in the *Xl*64 sample obtained from *E. coli* transformed with the FO synthase-producing plasmid, at an 8-HDF/FAD molar ratio similar to that observed for the in vitro reconstituted apo*Xl*64•HDF complex, whereas no 8-HDF was detected for *At*64 obtained by the co-expression system. These in vitro and in vivo reconstruction experiments revealed that (i) *Xenopus* (6–4)PL does bind 8-HDF while *Arabidopsis* (6–4)PL cannot, and (ii) the hydroxy group at the C8 position in 8-HDF participates in the 8-HDF uptake. It has been pointed out that the plant *At*64 would not utilize 8-HDF as the light-harvesting chromophore, because of lack of FO1 synthase homolog genes in the entire plant genome [18]. From the structural viewpoint, *At*64 shares the amino acid residues composing the 8-HDF binding site with *Dm*64 and *Cr*aCRY, but the crystal structure of *At*64 [35] suggests that the presence of an extended loop near the entrance of the binding site in *At*64 makes the entrance narrow (Figs. 4C and S1). In addition, Tyr117 in *At*64 is located at the position where the ribityl chain of 8-HDF exists in *Dm*64 and *Cr*aCRY (Fig. 4C), and thus the 8-HDF binding to *At*64 would be inhibited. In fact, the MD simulation of the *At*64 model indicated the POV of the putative 8-HDF binding pocket in *At*64 was 265 Å^3, (Fig. S3C), which is much smaller than that in apo*Dm*64 (502 Å^3) and rather close to the closed state of apo*Dm*64•HDF, although time development of their POVs exhibited similar fluctuation deviations (Fig. S3D). These results suggest that the putative 8-HDF binding pocket of *At*64 is not optimized for the 8-HDF uptake. However, there was a possibility that *At*64 could bind 8-HDF during the protein folding. Our in vivo reconstitution experiments clearly demonstrated that 8-HDF uptake into *At*64 did not occur even during the protein folding, suggesting that *At*64 would utilize different types of the second chromophore. Otherwise, *At*64 may not require the second chromophore for its function.

The above experiment using 8-MDF suggested a role of the 8-hydroxy group in the 8-HDF binding. Absorption spectra of 8-HDF at various pH revealed that the pK_a value of the 8-hydroxy group is estimated to be 6.1 ± 0.1 (Fig. S5), indicating that the 8-hydroxy group of 8-HDF exists in a deprotonated form (-O$^-$) in the neutral solution, with an absorption maximum at 422 nm ($\varepsilon_{422}=23{,}300\ \text{M}^{-1}\ \text{cm}^{-1}$). In its proximity in holo*Dm*64 and apo*Dm*64•HDF, there are two positively charged residues Arg60 and Lys266, which correspond to Arg51 and Lys256 in *Xl*64, respectively, and therefore these residues seem to be responsible for recognition of the phenoxide moiety of 8-HDF through

Coulombic interaction. In our POVME analysis (Fig. 3), Ile50 of apo*Dm*64●HDF was found to play a key role in the loop translocation. In *Xl*64, this residue is replaced with Phe41, which is the identical residue in *Cra*CRY (Fig. 2A). Hence, this Phe41 expectedly plays a similar role in the 8-HDF uptake. To identify the role of these residues in the 8-HDF binding, F41A, R51A, and K256A mutants of *Xl*64 were prepared, and the above experiments were performed (Fig. 4B). Interestingly, the F41A mutant exhibited a significantly reduced 8-HDF/FAD molar ratio as compared to the wild type, presumably due to the incomplete loop translocation as observed for our MD simulations. On the other hand, the single mutants, i.e., R51A and K256A, bound 8-HDF at the 8-HDF/FAD molar ratio similar to that obtained for the wild type. The R51A/K256A double mutant was also prepared, and the result indicated that the 8-HDF binding to the double mutant was slightly decreased, although it was out of the defined significance ($P = 0.076$). The results suggest that both of the two positively charged side chains are involved in the 8-HDF binding.

To analyze the 8-HDF binding in detail, *Xl*64 proteins were titrated into the 8-HDF solution, and the absorption changes were measured (Fig. 5A). Upon titration of *Xl*64, the absorption spectra of 8-HDF gradually changed, with a ~20 nm shift of absorption maximum wavelength (λ_{max}) from 422 to 441 nm. Conversely, titration of *At*64, which did not bind 8-HDF (Fig. 4B), into the 8-HDF solution simply yielded an increase of the FAD$_{ox}$ absorption without any spectral shift (Fig. S6). Altogether, the red shift of λ_{max} observed for *Xl*64 can be attributed to the change in the environment of the 8-HDF molecule, namely uptake from the polar aqueous solution into the apolar protein-binding pocket, as shown in Fig. 2B. As mentioned above, the mutations introduced in *Xl*64 apparently affected the 8-HDF binding (Fig. 4A), and thus the titration of the mutants into the 8-HDF solution was also investigated. The λ_{max} values of the single mutant samples (F41A, R51A and K256A) were shifted to ~440 nm (Fig. S6), suggesting the 8-HDF binding to these mutants. On the other hand, interestingly, titration of the R51A/K256A mutant into 8-HDF exhibited a distinct trend in the spectral change (Fig. 5B). After addition of the protein, two major peaks at 410 and 445 nm became prominent, the latter of which would partially be caused by the FAD$_{ox}$ absorption. Except this double mutant, λ_{max} was red-shifted upon titration, indicating the binding mode of the R51A/K256A mutant would be remarkably different from that of the others.

To extract the 8-HDF absorption in the bound state, the ideal FAD absorption spectra of *Xl*64, used in the previous study [26], were manually subtracted from the spectra after the final titration. Using the absorbance of 8-HDF of each measurement in the absence of the proteins as an internal standard, the estimated absorption spectra of 8-HDF bound in the respective proteins was normalized (Fig. 5C). In the case of *Xl*64, the 8-HDF in the bound state had 1.4-fold stronger absorption than in the isolated state at the neutral pH, with λ_{max} of 440 nm. The hyperchromic effect upon the binding was also observed for the R51A and K256A mutants, whereas the hyperchromic effect of R51A was slightly smaller than that of the wild type and the K256A mutant. On the other hand, the absorption spectrum of 8-HDF bound to the R51A/K256A double mutant exhibited a non-isomorphic structure with two peaks at 409 and 443 nm, indicating formation of two, or even more, different
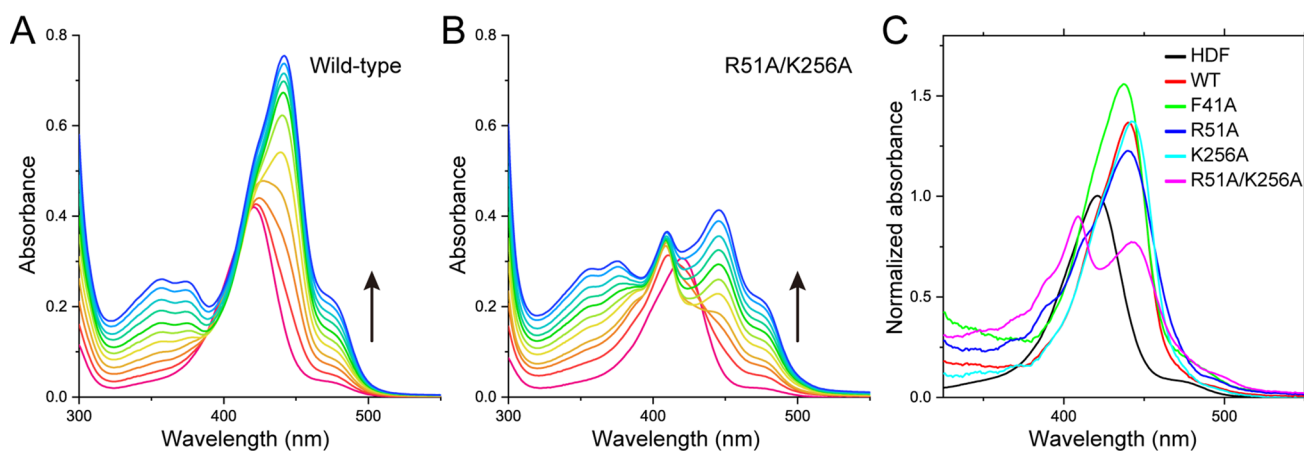

Fig. 5 Spectral changes upon titration of the *Xl*64 proteins into the 8-HDF solution. Increasing amounts (0–30 μM, red to blue) of *Xl*64 (**a**) and R51A/K256A (**b**) were titrated into 8-HDF. Spectral changes upon titration of *Xl*64 into 8-MDF and titration of F41A, R51A, K256A, and *At*64 into 8-HDF are shown in Fig. S6. **c** Normalized absorption spectra of 8-HDF in the isolated (black) and bound states (colored). Using the ideal FAD spectra used in the previous study, the absorption of FAD in the protein was subtracted from the last absorption spectrum of each experiment, so as to extract the absorption spectrum of 8-HDF bound to the respective protein. The obtained absorption spectra were normalized by setting the absorbance of 8-HDF at 422 nm in the isolated state in each experiment to 1

states. Judged from the 8-HDF spectrum bound to the wild-type protein, the peak at ~440 nm would correspond to the bound state. Our MM-PBSA calculation suggested that Arg60 in *Dm*64 (corresponding to Arg51 in *Xl*64) showed the stronger interaction with 8-HDF than Lys266 (corresponding to Lys256 in *Xl*64), and thus the spectrum of 8-HDF bound to the K256A mutant was in nice agreement with that bound to the wild type, while the reduced hyperchromic effect was observed for the R51A mutant. In the absence of Arg51, the main counterpart of the phenoxide moiety of 8-HDF, Lys256 may compensate and function to accommodate the 8-HDF molecule at the proper position, resulting in the similar absorption spectrum of 8-HDF in the R51A mutant. In the R51A/K256A double mutant, lack of two positively charged residues proximal to the phenoxide moiety caused a misdirection of 8-HDF to the inappropriate position in the binding pocket, presumably leading to the formation of blue-shifted species. Comprehensively, the Coulombic interactions between Arg51/Lys256 and 8-HDF plays a key role in the accommodation of 8-HDF.

The 8-HDF absorption spectrum in the F41A-bound form showed a broader band than in the other proteins, with λ_{max} of 435 nm, as if the absorption band consists of the mixture of the unbound and bound state of 8-HDF. To check the possibility, the unbound 8-HDF was removed from the mixture of 8-HDF and the proteins using the gel filtration column as mentioned above, and the absorption spectra of the eluted proteins were measured. Supposing that the 8-HDF binding process is in equilibrium, removal of the unbound 8-HDF from the mixture should induce the dissociation of 8-HDF as time goes by, and time-dependent blue-shift of λ_{max} could be observed. The results (Fig. S7) indicated that λ_{max} of the wild type, F41A, R51A, and K256A remained constant during 2 h after the gel filtration. These results suggest that the 8-HDF binding process is a one-way reaction. Our MD simulation of apo*Dm*64-I50A●HDF indicated the open and closed states of the 8-HDF binding site. Phe41 in *Xl*64 works as well as Ile50 in *Dm*64, and 8-HDF in the F41A mutant was partially exposed to the aqueous environment due to the incomplete loop translocation, resulting in the observation of the blue-shifted 8-HDF absorption.

3.3 DNA repair of Xl64 proteins bearing 8-HDF

From our analyses on 8-HDF binding to the *Xl*64 proteins, we could conclude that mutations of the amino acid residues did not strikingly impact the 8-HDF uptake, although how well 8-HDF is accommodated in the binding site was affected by the mutations. To evaluate to what extent the difference in the binding manner of 8-HDF impact the enzymatic function, DNA repair experiments using the 8-HDF bound *Xl*64 proteins were performed. To link the binding manner to the repair activity, the 8-HDF bound enzymes

reconstituted in vitro were exclusively used. Repair of the (6–4)PP upon white light illumination was monitored by HPLC (Fig. S8), and the number of repaired substrates per enzyme was plotted against the illumination time (Fig. 6). For comparison, the DNA repair activity of *Xl*64 without 8-HDF was also investigated in the same way (Fig. S8), and its linear regression is shown in Fig. 6 as a reference. *Xl*64●HDF exhibited 9-times higher DNA repair activity than the apoenzyme (see Table S2), suggesting that 8-HDF properly functions as the LHC for *Xl*64.

Regarding the mutants prepared for this study, their DNA repair abilities in the presence of 8-HDF were lower than the wild-type protein. Since the fluorescence spectra of 8-HDF in the wild type and mutants (F41A and R51A/K256A) of *Xl*64 were essentially identical to each other (Fig. S9), the reduced DNA repair activities of the mutant proteins bearing 8-HDF would be caused by disordered electronic coupling between the donor and acceptor. Among them, R51A and K256A mutants exhibited 50–60% activities as compared to *Xl*64●HDF. Interestingly, loss of the two residues significantly reduced the DNA repair activity to the level of the wild type without 8-HDF. As discussed above, these residues interact with the phenoxide moiety of 8-HDF through
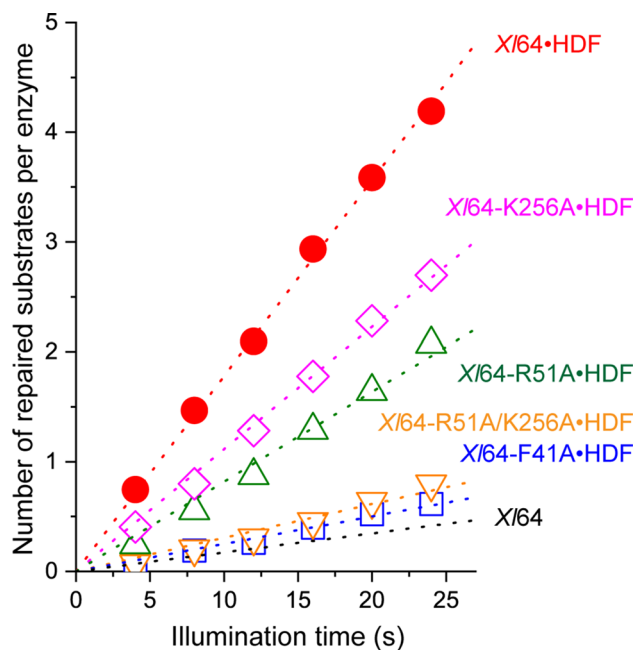

Fig. 6 Repair of 6–4PP by *Xl*64 proteins bearing 8-HDF. Number of repaired substrates per enzyme for *Xl*64●HDF (red filled circle), *Xl*64-F41A●HDF (blue open square), *Xl*64-R51A●HDF (green open triangle), *Xl*64-K256A●HDF (magenta open diamond), and *Xl*64-R51A/K256A●HDF (orange open inverted triangle) are plotted against illumination time. Their linear regressions are also shown in dashed lines. The black dashed line represents the linear regression of the DNA repair by *Xl*64 without 8-HDF, the original data of which are shown in Fig. S8. The slopes of the linear regressions are summarized in Table S2

Coulombic interactions. Without these residues, the 8-HDF cannot be accommodated in the binding site in the proper direction, and the energy transfer from the bound 8-HDF to FADH⁻ would barely take place, leading to the DNA repair activity similar to apoXl64. The F41A mutation also significantly reduced the DNA repair activity down to the level of apoXl64. Considering the role of Phe41 in the loop translocation of the binding site from the open to closed states, the functional analysis indicated that the translocation is also remarkably important for the energy transfer. As observed in Fig. 5C, the electronic state of 8-HDF in F41A was different from that in wild type, and the difference was caused by the surrounding circumstance of the fluorophore. The loop translocation, therefore, provides an apolar circumstance of 8-HDF for the efficient energy transfer, by inhibiting water molecules from coming close to 8-HDF.

These functional analyses in combination with the theoretical calculations suggest general requisites for accommodation of the functionally active LHC in Xl64 (Fig. 7). Since the orientation and distance between donor and acceptor critically affect the FRET kinetics, fine-tuning of the position and the electronic state of LHC is considerably important. Since the mutations in the 8-HDF binding site tested in this study did not cause complete loss of the 8-HDF uptake (Fig. 4B), the primary driving force for the 8-HDF uptake is supposed to be van der Waals interactions between the apolar amino acid side chains and the π-plane of the deazaisoalloxazine moiety. Upon the uptake, the phenoxide moiety is recognized by mainly Arg51, with a support of Lys256, through Coulombic interactions, resulting in the accommodation of 8-HDF in the proper direction in the binding site. After the 8-HDF uptake, the loop translocation induced by the through-space interaction (π–π or CH–π) between Phe41 in Xl64 or Ile50 in Dm64 and the deazaisoalloxazine ring takes place to close the entrance of the binding site, to ensure the apolar circumstance of 8-HDF. Noteworthy, the amino acid residues important for the 8-HDF uptake and function
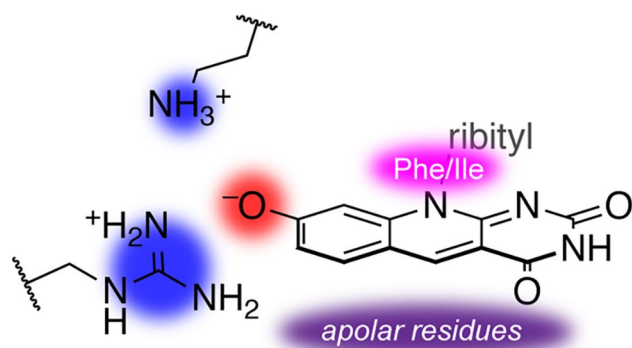
are well conserved in *Anacystis nidulans* CPD photolyase, which is also known as an 8-HDF binding protein [23] (Fig. S10). On the other hand, proteins in the photolyase/crypto-chome superfamily that utilize MTHF or FMN as their LHC tend to lack the hydrophobic translocating Phe/Ile residue and the positively-charged Lys residues responsible for the recognition of ribityl and phenoxide moieties in 8-HDF (Fig. S10). Hence, these residues may be useful to predict the LHC of this superfamily.

3.4 Conclusions and perspectives

Our comprehensive study demonstrated how *Xenopus* (6–4) PL incorporates 8-HDF in the binding site and utilizes the molecule as the light-harvesting chromophore. The insights will provide a guideline for designing artificial LHCs that potentially enhance DNA repair activity of (6–4)PL. The previous study showed that modification of *Thermus thermophilus* DNA photolyase with the synthetic coumarin chromophore enhances its DNA repair activity [9]. In this case, the artificial LHC has been covalently attached to the enzyme, and one of other possibilities for the enhancement of the enzymatic activity of PLs is to bind artificial LHCs utilizing the natural LHC binding site. Functional modification of 8-HDF by keeping the interactions required for the uptake will thus lead to the development of an artificial light-harvesting complex of (6–4)PL.

The orientation factor and kinetics of the energy transfer between the LHC and FADH⁻ can be experimentally estimated by time-resolved fluorescent spectroscopy, and these parameters for the natural 8-HDF–FADH⁻ FRET pair should be estimated in future. As the amplitude and the direction of the transition dipole moment of the artificial LHCs can be foreseen by theoretical calculations, comparison of the obtained parameters with those of 8-HDF in addition to the extrapolation of the orientation factor of the 8-HDF–FADH⁻ pair will enable us to predict the function of designed molecules, when the binding manner is conserved.

Author contributions J.Y. conceived the idea and directed the research. Y.H., H.M., R.S., and J.Y. performed the MD simulations. A.M. and R.K.V. synthesized the 8-HDF derivatives. A.M. and Y.H. prepared

Fig. 7 Plausible interactions required for the accommodation of functionally active 8-HDF

proteins for in vitro reconstitution. Y.H. and J.Y. prepared proteins for in vivo reconstitution. A.M. performed the 8-HDF binding and DNA repair experiments. Y.H., S.I., R.S., and J.Y. analyzed the data and wrote the manuscript, and all the authors reviewed the manuscript.

Declarations

References

1. Shen, J.-R. (2015). The structure of photosystem II and the mechanism of water oxidation in photosynthesis. *Annual Review of Plant Biology, 66*, 23–48. https://doi.org/10.1146/annurev-arplant-050312-120129

2. Morris, M. B., Dastmalchi, S., & Church, W. B. (2009). Rhodopsin: structure, signal transduction and oligomerization. *The International Journal of Biochemistry and Cell Biology, 41*(4), 721–724. https://doi.org/10.1016/j.biocel.2008.04.025

3. van der Horst, M. A., Laan, W., Yeremenko, S., Wende, A., Palm, P., Oesterhelt, D., & Hellingwerf, K. J. (2005). From primary photochemistry to biological function in the blue-light photoreceptors PYP and AppA. *Photochemical and Photobiological Sciences, 4*(9), 688–693. https://doi.org/10.1039/B418442B

4. Sancar, A. (2003). Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chemical Reviews, 103*(6), 2203–2237. https://doi.org/10.1021/cr0204348

5. Brettel, K., & Byrdin, M. (2010). Reaction mechanisms of DNA photolyase. *Current Opinion in Structural Biology, 20*(6), 693–701. https://doi.org/10.1016/j.sbi.2010.07.003

6. Croce, R., & van Amerongen, H. (2014). Natural strategies for photosynthetic light harvesting. *Nature Chemical Biology, 10*(7), 493–501. https://doi.org/10.1038/nchembio.1555

7. Springer, J. W., Parkes-Loach, P. S., Reddy, K. R., Krayer, M., Jiao, J., Lee, G. M., Niedzwiedzki, D. M., Harris, M. A., Kirmaier, C., Bocian, D. F., Lindsey, J. S., Holten, D., & Loach, P. A. (2012). Biohybrid photosynthetic antenna complexes for enhanced light-harvesting. *Journal of the American Chemical Society, 134*(10), 4589–4599. https://doi.org/10.1021/ja207390y

8. Nagakawa, H., Takeuchi, A., Takekuma, Y., Noji, T., Kawakami, K., Kamiya, N., Nango, M., Furukawa, R., & Nagata, M. (2019). Efficient hydrogen production using photosystem I enhanced by artificial light harvesting dye. *Photochemical and Photobiological Sciences, 18*(2), 309–313. https://doi.org/10.1039/C8PP00426A

9. Terai, Y., Sato, R., Matsumura, R., Iwai, S., & Yamamoto, J. (2020). Enhanced DNA repair by DNA photolyase bearing an artificial light-harvesting chromophore. *Nucleic Acids Research, 48*(18), 10076–10086. https://doi.org/10.1093/nar/gkaa719

10. Kim, S. T., Heelis, P. F., Okamura, T., Hirata, Y., Mataga, N., & Sancar, A. (1991). Determination of rates and yields of interchromophore (folate→flavin) energy transfer and intermolecular (flavin→DNA) electron transfer in Escherichia coli photolyase by time-resolved fluorescence and absorption spectroscopy. *Biochemistry, 30*(47), 11262–11270. https://doi.org/10.1021/bi00111a011

11. Cong, H., Niedzwiedzki, D. M., Gibson, G. N., LaFountain, A. M., Kelsh, R. M., Gardiner, A. T., Cogdell, R. J., & Frank, H. A. (2008). Ultrafast time-resolved carotenoid to bacteriochlorophyll energy transfer in LH2 complexes from photosynthetic bacteria. *Journal of Physical Chemistry B, 112*(34), 10689–10703. https://doi.org/10.1021/jp711946w

12. Visser, A. J. W. G., Laptenok, S. P., Visser, N. V., van Hoek, A., Birch, D. J. S., Brochon, J.-C., & Borst, J. W. (2010). Time-resolved FRET fluorescence spectroscopy of visible fluorescent protein pairs. *European Biophysics Journal, 39*(2), 241–253. https://doi.org/10.1007/s00249-009-0528-8

13. Tan, C., Guo, L., Ai, Y., Li, J., Wang, L., Sancar, A., Luo, Y., & Zhong, D. (2014). Direct determination of resonance energy transfer in photolyase: structural alignment for the functional state. *Journal of Physical Chemistry A, 118*(45), 10522–10530. https://doi.org/10.1021/jp504349b

14. Yoneda, Y., Noji, T., Katayama, T., Mizutani, N., Komori, D., Nango, M., Miyasaka, H., Itoh, S., Nagasawa, Y., & Dewa, T. (2015). Extension of light-harvesting ability of photosynthetic light-harvesting complex 2 (LH2) through ultrafast energy transfer from covalently attached artificial chromophores. *Journal of the American Chemical Society, 137*(40), 13121–13129. https://doi.org/10.1021/jacs.5b08508

15. Scholes, G. D. (2003). Long-range resonance energy transfer in molecular systems. *Annual Review of Physical Chemistry, 54*, 57–87. https://doi.org/10.1146/annurev.physchem.54.011002.103746

16. Yamamoto, J., Plaza, P., & Brettel, K. (2017). Repair of (6–4) lesions in DNA by (6–4) photolyase: 20 years of quest for the photoreaction mechanism. *Photochemistry and Photobiology, 93*(1), 51–66. https://doi.org/10.1111/php.12696

17. Park, H.-W., Kim, S.-T., Sancar, A., & Deisenhofer, J. (1995). Crystal structure of DNA photolyase from *Escherichia coli*. *Science, 268*(5219), 1866–1872. https://doi.org/10.1126/science.7604260

18. Franz, S., Ignatz, E., Wenzel, S., Zielosko, H., Ngurah Putu, E. P. G., Maestre-Reyna, M., Tsai, M.-D., Yamamoto, J., Mittag, M., & Essen, L.-O. (2018). Structure of the bifunctional cryptochrome aCRY from *Chlamydomonas reinhardtii*. *Nucleic Acids Research, 46*(15), 8010–8022. https://doi.org/10.1093/nar/gky621

19. Zhang, F., Scheerer, P., Oberpichler, I., Lamparter, T., & Krauß, N. (2013). Crystal structure of a prokaryotic (6–4) photolyase with an Fe-S cluster and a 6,7-dimethyl-8-ribityllumazine antenna chromophore. *Proceedings of the National Academy of Sciences of the United States of America, 110*(18), 7217–7222. https://doi.org/10.1073/pnas.1302377110

20. Fujihashi, M., Numoto, N., Kobayashi, Y., Mizushima, A., Tsujimura, M., Nakamura, A., Kawarabayasi, Y., & Miki, K. (2007). Crystal structure of archaeal photolyase from *Sulfolobus tokodaii* with two FAD molecules: implication of a novel light-harvesting cofactor. *Journal of Molecular Biology, 365*(4), 903–910. https://doi.org/10.1016/j.jmb.2006.10.012

21. Ueda, T., Kato, A., Kuramitsu, S., Terasawa, H., & Shimada, I. (2005). Identification and characterization of a second chromophore of DNA photolyase from Thermus thermophilus HB27. *Journal of Biological Chemistry, 280*(43), 36237–36243. https://doi.org/10.1074/jbc.M507972200

22. Kiontke, S., Gnau, P., Haselsberger, R., Batschauer, A., & Essen, L.-O. (2014). Structural and evolutionary aspects of antenna chromophore usage by class II photolyases. *Journal of Biological*

Chemistry, 289(28), 19659–19669. https://doi.org/10.1074/jbc.M113.542431

23. Eker, A. P. M., Kooiman, P., Hessels, J. K. C., & Yasui, A. (1990). DNA photoreactivating enzyme from the cyanobacterium Anacystis nidulans. *Journal of Biological Chemistry, 265*(14), 8009–8015. https://doi.org/10.1016/S0021-9258(19)39031-3

24. Glas, A. F., Maul, M. J., Cryle, M., Barends, T. R. M., Schneider, S., Kaya, E., Schlichting, I., & Carell, T. (2009). The archaeal cofactor F_0 is a light-harvesting antenna chromophore in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America, 106*(28), 11540–11545. https://doi.org/10.1073/pnas.0812665106

25. Hossain, M. S., Le, C. Q., Joseph, E., Nguyen, T. Q., Johnson-Winters, K., & Foss, F. W. (2015). Convenient synthesis of deazaflavin cofactor FO and its activity in F_{420}-dependent NADP reductase. *Organic & Biomolecular Chemistry, 13*(18), 5082–5085. https://doi.org/10.1039/c5ob00365b

26. Yamamoto, J., Shimizu, K., Kanda, T., Hosokawa, Y., Iwai, S., Plaza, P., & Müller, P. (2017). Loss of fourth electron-transferring tryptophan in animal (6–4) photolyase impairs DNA repair activity in bacterial cells. *Biochemistry, 56*(40), 5356–5364. https://doi.org/10.1021/acs.biochem.7b00366

27. Hosokawa, Y., Sato, R., Iwai, S., & Yamamoto, J. (2019). Implications of a water molecule for photoactivation of plant (6–4) photolyase. *Journal of Physical Chemistry B, 123*(24), 5059–5068. https://doi.org/10.1021/acs.jpcb.9b03030

28. Maul, M. J., Barends, T. R. M., Glas, A. F., Cryle, M. J., Domratcheva, T., Schneider, S., Schlichting, I., & Carell, T. (2008). Crystal structure and mechanism of a DNA (6–4) photolyase. *Angewandte Chemie International Edition, 47*(52), 10076–10080. https://doi.org/10.1002/anie.200804268

29. Case, D., Cerutti, D., Cheateham, T., Darden, T., Duke, R., Giese, T., Gohlke, H., Goetz, A., Greene, D., Homeyer, N., Simmerling, C., Botello-Smith, W., Swail, J., Walker, R., Wang, J., Wolf, R., Wu, X., Xiao, L., & Kollman, P. (2016). *Amber 2016*. San Francisco: University of California.

30. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation, 11*(8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255

31. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics, 79*, 926–935. https://doi.org/10.1063/1.445869

32. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd,

J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., Fox, D. J. (2016). Gaussian16, Revision C.01, Gaussian, Inc., Wallingford CT, 2016.

33. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry, 25*(9), 1157–1174. https://doi.org/10.1002/jcc.20035

34. Sousa da Silva, A. W., & Vranken, W. F. (2012). ACPYPE—Ante-Chamber PYthon parser interfacE. *BMC Research Notes, 5*, 367. https://doi.org/10.1186/1756-0500-5-367

35. Hitomi, K., DiTacchio, L., Arvai, A. S., Yamamoto, J., Kim, S.-T., Todo, T., Tainer, J. A., Iwai, S., Panda, S., & Getzoff, E. D. (2009). Functional motifs in the (6–4) photolyase crystal structure make a comparative framework for DNA repair photolyases and clock cryptochromes. *Proceedings of the National Academy of Sciences of the United States of America, 106*(17), 6962–6967. https://doi.org/10.1073/pnas.0809180106

36. Abraham, M. J., Murtola, T., Schulz, R., Pall, Z., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

37. Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics, 81*(1), 511–519. https://doi.org/10.1063/1.447334

38. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics, 81*(8), 3684–3690. https://doi.org/10.1063/1.448118

39. Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry, 18*(12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12%3c1463::AID-JCC4%3e3.0.CO;2-H

40. Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics, 98*(12), 10089–10092. https://doi.org/10.1063/1.464397

41. Kumari, R., Kumar, R., Open Source Drug Discovery Consortium, & Lynn, A. (2014). g_mmpbsa–A GROMACS tool for high-throughput MM-PBSA calculations. *Journal of Chemical Information and Modeling, 54*(7), 1951–1962. https://doi.org/10.1021/ci500020m

42. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America, 98*(18), 10037–10041. https://doi.org/10.1073/pnas.181342398

43. Eisenberg, D., & McLachlan, A. D. (1986). Solvation energy in protein folding and binding. *Nature, 319*(6050), 199–203. https://doi.org/10.1038/319199a0