Enhanced DNA repair by DNA photolyase bearing an artificial light-harvesting chromophore

Yuma Terai[1], Ryuma Sato[2], Risa Matsumura[1], Shigenori Iwai[1] and Junpei Yamamoto [1],*

[1]Division of Chemistry, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan and [2]Center for Biosystems Dynamics Research, RIKEN, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

ABSTRACT

Photolyases are flavoenzymes responsible for the repair of carcinogenic DNA damage caused by ultraviolet radiation. They harbor the catalytic cofactor flavin adenine dinucleotide (FAD). The light-driven electron transfer from the excited state of the fully-reduced form of FAD to the DNA lesions causes rearrangement of the covalent bonds, leading to the restoration of intact nucleobases. In addition to the catalytic chromophore, some photolyases bear a secondary chromophore with better light absorption capability than FAD, acting as a light-harvesting chromophore that harvests photons in sunlight efficiently and transfers light energy to the catalytic center, as observed in natural photoreceptor proteins. Inspired by nature, we covalently and site-specifically attached a synthetic chromophore to the surface of photolyase using oligonucleotides containing a modified nucleoside and a cyclobutane-type DNA lesion, and successfully enhanced its enzymatic activity in the light-driven DNA repair. Peptide mapping in combination with theoretical calculations identified the amino acid residue that binds to the chromophore, working as an artificial light-harvesting chromophore. Our results broaden the strategies for protein engineering and provide a guideline for tuning of the light perception abilities and enzymatic activity of the photoreceptor proteins.

GRAPHICAL ABSTRACT
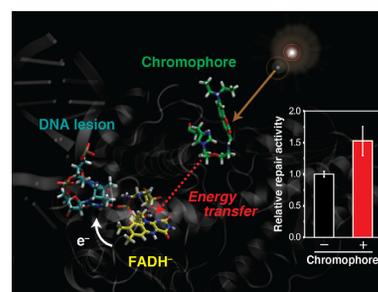
INTRODUCTION

Photolyases (PLs) are flavoproteins responsible for repair of carcinogenic UV-induced DNA damage (i.e. cyclobutane pyrimidine dimers (CPDs) and pyrimidine(6–4)pyrimidone photoproducts (6–4PPs)) by using blue light in sunlight (1). PLs exist in all domain of life except mammals (1), thus they have played a key role in the maintenance of genetic integrity against UV since the very beginning of the Earth history. The catalytic DNA repair by DNA photolyases requires the light-driven electron transfer from the excited state of fully reduced flavin adenine dinucleotide ($FADH^-$) to the lesions (Figure 1A), inducing the structural rearrangement to the intact nucleobases within a few nano-seconds for CPD repair (2,3). The absorption band of $FADH^-$ is distributed in the range of 350–500 nm, and one can estimate that $FADH^-$ perceives 0.4 photons s^{-1} per molecule by solar light (4). Considering the reported quantum yield of the CPD repair by PL (50–100%) (5–8), the CPD repair by PL proceeds in a highly efficient manner *in vitro*. Even *in vivo*, expression of the PL gene in transgenic mice reportedly increased the tolerance against UV (9), thus implying PLs work as a UV protectant.

Although the catalytic reaction of PLs exclusively relies on the $FADH^-$ excitation by blue light, living tissues are poorly permeable to blue light. To gain the photoreception ability, some PLs utilize secondary auxiliary chromophores that absorb blue light much more strongly than $FADH^-$. The excitation energy of the sec-

*To whom correspondence should be addressed. Tel: +81 6 6850 6240; Fax: +81 6 6850 6240; Email: yamamoto@chem.es.osaka-u.ac.jp
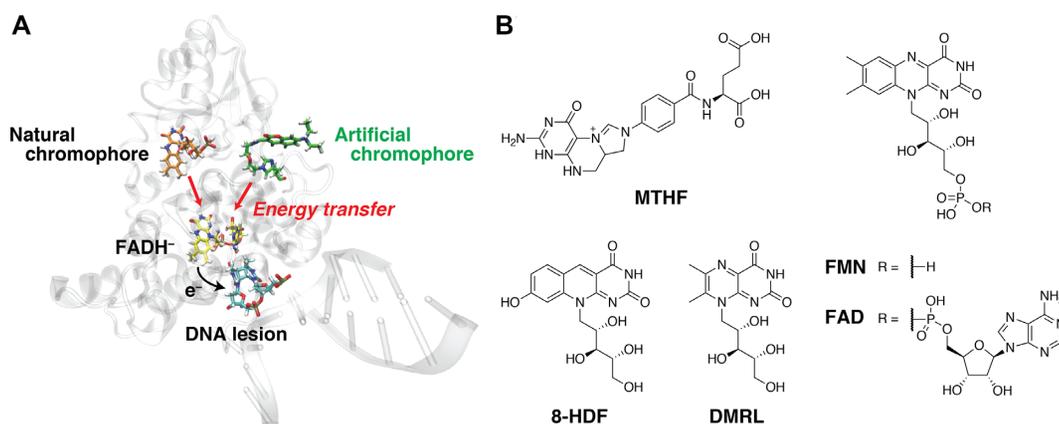
Figure 1. Antenna chromophores of DNA photolyases. (**A**) Schematic illustration of the light-driven DNA repair of PL enhanced by the energy transfer from a natural or an artificial chromophore to FADH⁻ via Förster resonance energy transfer. The model structure was created with the CPD photolyase from *Thermus thermophilus* (PDB ID: 2J09 (22)). The natural light-harvesting chromophore is embedded in the second chromophore binding pocket, whereas the artificial chromophore conjugated to an amino acid residue is exposed to the solvent. (**B**) Five natural antenna chromophores in PLs. These molecules commonly absorb blue light much more strongly than FADH⁻.

ondary chromophores is transferred to the ground state of FADH⁻, leading to the enhanced DNA repair by PLs. As the secondary chromophores harvest photon energy for the catalytic reaction, they are called as light-harvesting chromophores or antenna chromophores (Figure 1A). Contrary to the exclusively conserved FAD chromophore in PLs, the antenna chromophores in PLs are diverse. So far, five natural antenna chromophores with a pteridine architecture have been reported for PLs (Figure 1B): (i) 5,10-methenyltetrahydrofolate (MTHF) for bacterial (10,11) and algal PLs (12), (ii) flavin mononucleotide (FMN) (13) or (iii) FAD (14) for thermophilic bacterial PLs, (iv) 8-hydroxy-7,8-didemethyl-5-deazariboflavin (8-HDF) for class II CPD-repairing PLs (CPD-PL) (15), *Synechococcus elongatus* (also known as *Anacystis nidulans*) PL (16), 6–4PP-repairing PLs ((6–4)PLs) (17), and (v) 6,7-dimethyl-8-ribityllumazine (DMRL) for bacterial (6–4)PL (18). These molecules are well recognized in the binding site at the cleft between the N-terminal α/β domain and the catalytic C-terminal domain in PLs (Figure 1A) and emit fluorescence that overlaps with the absorption of FADH⁻, allowing them to transfer the energy presumably via the Förster mechanism (Förster resonance energy transfer, FRET) (19–21). Although such energy transfer using antenna chromophores is elegantly utilized in the photosynthesis, these antenna chromophores in PLs, the fluorescence of which overlaps with the red edge of FADH⁻ absorption spectrum, function as a supplement of FADH⁻.

One can, therefore, expect that tuning of the spectral properties of the antenna chromophore would afford the modulation of the photoreception ability and enzymatic activity of PL. Such a spectral tuning of PL using artificial flavin derivatives has been investigated in a crystallographic study (22), by soaking crystals of *Thermus thermophilus* CPD-PL (*Tth*CPD) into solutions of the flavin derivatives. Among the tested derivatives, however, only 8-iodoflavin was obviously incorporated into the binding site, but its spectral tuning ability was not tested. Since that report, to our best knowledge, no successful trials for the tuning of PL activity using artificial molecules have been re-

ported, presumably due to the difficulties in the uptake of artificial molecules into the binding site. An alternative way to introduce artificial molecules into proteins is covalent attachment of a small molecule containing a reactive moiety, such as the *N*-hydroxysuccinimide ester, to the amino acid side chain of lysine in this case (23). The drawback in the modification of proteins using small molecules is the uncontrollable labeling positions and the numbers of labeled residues, in addition to the unexpected occupation of the active site due to the labeling reaction, which would lead to inactivation of the enzyme.

Here, we report a ligand-directed labeling of PL with an artificial coumarin chromophore, the fluorescence spectrum of which overlaps with the FADH⁻ absorption. This strategy allowed us to modify PL chemically with a fluorophore and to produce the active and labeled PLs. Remarkably, one of the labeled PLs exhibited an enhanced light-driven DNA repair activity, indicating the synthetic fluorophore functions as a light harvesting chromophore. Peptide mapping and theoretical calculations identified the amino acid residue that binds to the chromophore, which is functional in the energy transfer. This work, thus, provides the first step for the tuning and modulation of the enzymatic activities of PLs using various types of fluorophores.

MATERIALS AND METHODS

General

Reagents for DNA synthesis were purchased from Glen Research. All other reagents and solvents were purchased from FUJIFILM Wako Pure Chemical Corporation, Sigma-Aldrich, Tokyo Chemical Industry, or Thermo Fisher Scientific. HPLC analyses were performed on a Gilson gradient-type analytical system equipped with a Waters 2998 photodiode-array detector. DNA samples were analyzed and purified by using a Waters μBondasphere C18 5 μm 300Å column (3.9 mm × 150 mm), 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) generated over 20 min. Peptide samples were analyzed by using

a GL Sciences Inertsil ODS-3 5 μm column (4.6 mm × 250 mm), at a flow rate of 1.0 ml min^{-1} with a linear gradient of acetonitrile in 0.1% trifluoroacetate (TFA) generated over 45 min. Absorption and fluorescent spectra were measured on a PerkinElmer Lambda 35 UV/Vis spectrometer and a JASCO FP-6500 spectrofluorometer, respectively. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF MS) was performed with a Bruker ultraflex III using the negative mode for oligonucleotides. MALDI TOF MS and MS/MS analyses of peptides were performed on the same device using the positive mode.

Synthesis of protein-labeling oligonucleotides

Oligonucleotides containing both an amino linker and the *cis-syn* CPD were synthesized on an Applied Biosystems 3400 DNA synthesizer, using the phosphoramidite building blocks of amino modifier C2-dT (Glen Research), *cis–syn* CPD with the natural internucleotide linkage (synthesized as previously described (24)), and protected deoxyribonucleosides for ultra-mild DNA synthesis (Glen Research), which were assembled on the solid support. Phenoxyacetic anhydride, tetrazole and iodine were used as capping, activation, and oxidation reagents, respectively. The synthesized oligonucleotides were cleaved from the support by treatment with 28% ammonia water at room temperature for 1 h and deprotected by heating at 55°C for 4 h. The products were purified by HPLC with an acetonitrile gradient from 6% to 13%. The purified oligonucleotides (40 nmol) were dissolved in 28.8 μl of 100 mM carbonate/bicarbonate buffer (pH 8.5), and then 3.2 μl of a 0.1 M DMSO solution of NHS-DEAC, which was prepared as described previously (25), was added to the oligonucleotide solutions. The mixtures were incubated at room temperature for 4 h. After ethanol precipitation, the obtained protein-labeling oligonucleotides were analyzed and purified by HPLC with an acetonitrile gradient from 9.5% to 32%. The purified oligonucleotides were analyzed by MALDI-TOF MS using recrystallized 3-hydroxypicolinic acid as a matrix (Supplementary Table S1). The amounts of these oligonucleotides were determined by the absorption of DEAC moiety using its molar extinction coefficient of 18 800 M^{-1} cm^{-1} at 428 nm.

Preparation of recombinant *T. thermophilus* CPD photolyase

The codon-optimized cDNA of *Tth*CPD was purchased from Genewiz and was subcloned into the NdeI/BamHI restriction enzyme site of the modified pET-16b vector, which carries a FLAG-tag 5′ upstream of the N-terminal His-tag. After the DNA sequencing of the obtained plasmid, *Escherichia coli* C41(DE3) cells were transformed with the plasmid and cultured in 4 l of Luria-Broth medium with 50 μg ml^{-1} ampicillin at 25°C until the turbidity measured at 600 nm reached 0.8. Then, isopropyl 1-thio-β-D-galactopyranoside was added to the medium at a final concentration of 0.1 mM. After incubation at 25°C for 24 h, the cells were harvested and resuspended in 40 ml of buffer A (50 mM phosphate, 200 mM NaCl, 5 mM imidazole, and 10% glycerol (pH 8)) plus lysozyme at a final concentration

of 1 mg ml^{-1}. The cells were disrupted by sonication, and the mixture was treated at 65°C for 15 min to remove endogenous proteins derived from *E. coli*. The mixture was centrifuged at 12 000 rpm for 1 h, and the supernatant was poured into an open column filled with the TALON metal affinity resin (Clontech) equilibrated with buffer A. The resin was washed with 4 column volumes (CVs) of buffer B (50 mM phosphate, 200 mM NaCl, 10 mM imidazole, 5% glycerol (pH 8)), and the His-tagged protein was eluted with 4 CVs of buffer C (50 mM phosphate, 200 mM NaCl, 500 mM imidazole, 5% glycerol (pH 8)). The deep green elute was loaded onto a HiTrap Heparin HP column (GE Healthcare). The column was washed with a stepwise gradient of 100–500 mM NaCl in a buffer containing 50 mM Tris-HCl and 5% glycerol (pH 8). The fractions eluted from the columns were analyzed by 10% SDS-PAGE, and the deep green colored fractions were concentrated with Amicon Ultra-4 centrifugal filter unit (Merck Millipore, molecular weight cutoff: 30 kDa). The amounts of the obtained proteins were quantified by the Bradford assay using bovine serum albumin as a standard.

Evaluation of the labeling yields by protein-labeling oligonucleotides

Purified *Tth*CPD (20 μM) was incubated with protein-labeling oligonucleotides (20 μM) in 60 μl of 100 mM phosphate buffer (pH 8) at 37°C. After 0, 1, 2, 4, 8 and 24 h, aliquots (10 μl) of the samples were mixed with an equal volume of 2× SDS-PAGE loading buffer, and then applied to 10% SDS-PAGE. After running at a constant voltage (150 V) for 1.5 h, the gel was analyzed by in-gel fluorescence imaging. The 2:1 DNA–photolyase sample was also run through the gel and used as an internal standard to calibrate the fluorescence intensities of each sample. After the fluorescence detection, the gel was stained with Coomassie Brilliant Blue (CBB). The gel images were obtained with a ChemiDoc MP imager (Bio-Rad). The in-gel fluorescence images were analyzed using Image Lab software (Bio-Rad), and the band intensities were quantified by subtracting the background from the detected band intensities. The experiments were independently performed in triplicate, and the mean values and standard deviations were plotted over the reaction times.

Recovery of the fluorescently-labeled *Tth*CPD

Purified *Tth*CPD (20 μM) was incubated with protein-labeling oligonucleotides (20 μM) in 600 μl of 100 mM phosphate buffer (pH 8) at 37°C. After 24 h, 15 μl of 200 mM L-cysteine was added to the solution to give a final concentration of 5 mM. The mixture in the test tube was illuminated with white light to dissociate the DNA–protein complexes. After 2 h of illumination, the protein was purified using the HiTrap Heparin HP column as above mentioned. The colored fractions were concentrated with Amicon Ultra-0.5 centrifugal filter unit (Merck Millipore, molecular weight cutoff: 30 kDa), and the concentration of the fluorescently-labeled *Tth*CPD was quantified by the absorption of semiquinoid flavin adenine dinucleotide (FADH•) using its molar extinction coefficient of 4700 M^{-1}

cm^{-1} at 579 nm. Notably, the concentration determined by the FADH• absorption was comparable with that obtained by the Bradford assay.

DNA repair experiments

The purified enzyme (2 μM) in reaction buffer (94 μl, 20 mM phosphate, 500 mM NaCl, and 10% glycerol (pH 7.5)) was placed in a 10 mm × 2 mm × 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 both 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 mm × 8 mm window, in order to photoreduce FAD to fully-reduced flavin adenine dinucleotide (FADH$^-$), the single-stranded 14-mer substrate (5′-ATCGGCTTCGCGCA-3′) 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 either white light (430–800 nm, 10 s) or monochromatic light (380 or 428 nm, 10 s) from a MAX-150 xenon lamp (Asahi Spectra) through the 10 mm × 8 mm window for a total illumination time of 60 s, with an interval resting time of 3 min. After each illumination, aliquots (10 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. The experiments were performed in triplicate, and the mean values and the standard deviations were plotted over either the irradiation times under the white light conditions or the photon fluences under the monochromatic light conditions. The data were fitted with a linear model, and the slope was defined as the repair activity. The relative repair activities of the tested enzymes are shown with bar graphs, using non-labeled *Tth*CPD (**PL**) as a standard.

Evaluation of FMN contents of tested *Tth*CPD in the repair experiments

Fluorescently-labeled and non-labeled *Tth*CPDs (340 pmol) were deproteinized with phenol/chloroform/isoamyl alcohol (25:24:1), and the supernatants were analyzed by HPLC with an acetonitrile gradient from 7% to 11% to investigate the presence of FAD and FMN. The FMN contents of tested *Tth*CPDs were calculated by stipulating the FAD content as the protein amount.

Determination of the photorepair quantum yield

From the data of the **PL** repair experiments using monochromatic light at 380 nm, the numbers of repaired substrates per enzyme were plotted over irradiation times, and the linear regression of the data indicated the reaction rate constant of 0.074 s^{-1}. According to a previously reported chemical actinometry method (5), *p*-(dimethylamino)-benzenediazonium tetrafluoroborate (DMAD) was used as a chemical actinometer to deter-

mine the photorepair quantum yield of *Tth*CPD. The sample solution (200 μl) consisting of 2 μM DMAD and 50 mM H$_2$SO$_4$ was placed in the same cuvette as described in the repair experiment section and illuminated on ice with monochromic light at 380 nm from the MAX-150 xenon lamp (Asahi Spectra) through the 10 mm × 8 mm window. After each illumination for 1 s, the absorption spectrum of the sample was measured through the 10 mm path on the UV/Vis spectrometer. The natural logarithms of the obtained absorbances at 378 nm were plotted over the illumination times, and the data were fitted with a linear model. By taking into account the rate constants of **PL**, the slope of the DMAD photoreaction, the molecular extinction coefficients of FADH$^-$ (5600 M^{-1} cm^{-1}) and DMAD at 380 nm (5), and the quantum yield of the chemical actinometer, we determined a photorepair quantum yield of 54%. In the case of the 428 nm illumination, we obtained a reaction rate constant of 0.038 s^{-1} from the regression of the data and an excitation rate of 0.067 s^{-1}, which was calculated using the molecular extinction coefficient of FADH$^-$ (2100 M^{-1} cm^{-1} at 428 nm) and the incident photon flux in the repair experiment. From the parameters, we also determined a photorepair quantum yield of 57% at 428 nm.

Peptide mapping of fluorescently-labeled *Tth*CPDs

Purified *Tth*CPD (20 μM) was incubated with modified oligonucleotides (20 μM) in 1.2 ml of 100 mM phosphate buffer (pH 8) at 37°C for 24 h. The mixture was buffer-exchanged to the denaturing buffer (50 mM Tris–HCl and 8M urea (pH 8)) using Amicon Ultra-0.5 centrifugal filter unit (Merck Millipore, molecular weight cut-off: 30 kDa), and 20 μl of 500 mM dithiothreitol was added to the solution to give a final concentration of 20 mM. After incubation at 95°C for 10 min, the resulting solution was again buffer-exchanged to 50 mM Tris–HCl buffer (pH 8). To this solution, 16 μl of trypsin dissolved in 50 mM acetic acid at a concentration of 1 μg μl^{-1} was added at a trypsin/photolyase ratio of 1/75 (w/w). After incubation overnight at 37°C, peptides were passed through Pierce C18 spin columns (Thermo Fisher Scientific). The fluorescently-labeled fragments were purified by HPLC with an acetonitrile gradient from 20% to 60% and analyzed by MALDI-TOF MS/MS using recrystallized 4-hydroxy-α-cyanocinnamic acid as a matrix.

Molecular dynamics simulations

To explore the behavior of *Tth*CPD modified with DEAC, we performed molecular dynamics (MD) simulations. As the amino acid residues of His97, His123, Asp128, and His257 were found to be labeled with DEAC in **PL$_{ODN1}$**, these residues in the *Tth*CPD structure (PDB ID: 2J09 (22)) were modified with DEAC through the linker. In the case of the histidine side chain (His97, His123 and His257), DEAC was added to the ε- and δ-N atoms (termed as type1 and type2, respectively). In the case of the aspartic acid side chain (Asp128), DEAC was added to either of the two oxygen atoms in the carboxy group (termed as type1 and type2). The initial structures of the calculation systems containing these modified amino acids are shown in Supplementary Figure S1.

To create the force field of each amino acid residue with the DEAC moiety, we performed the geometry optimization of the molecules using Gaussian16 (26) at the B3LYP/6–31G(d) level of theory and consecutively performed the electronic structure calculation at the HF/6–31+G(d) level of theory. The geometry optimization and the electronic structure calculation of FADH$^-$ and FMN were also performed at the B3LYP/6–31G(d) and MP2/6–311++G(2d,2p) levels of theory, respectively. The amber force field of each modified amino acid residue, FADH$^-$, and FMN were created from the results of the electronic structure calculation using the antechamber module of the AMBER16 program package (27). Then, we applied the amber force field ff14SB (28) and the TIP3P water model (29) for the protein and the water molecules, respectively, and the created force fields were also applied to each modified amino acid residue, FADH$^-$ and FMN, using the tLeap module of the AMBER16 program package. The systems were then neutralized with counter ions. The final systems consisted of 58 784 (His97-type1), 58 862 (His97-type2), 58 799 (His123-type1), 58 841 (His123-type2), 58 777 (Asp128-type1), 58 885 (Asp128-type2), 58 829 (His257-type1) and 58 883 (His257-type2) atoms including 17 281, 17 307, 17 286, 17 300, 17 278, 17 314, 17 296 and 17 314 water molecules and 3, 3, 3, 3, 4, 4, 3 and 3 chloride ions, respectively.

The AMBER topology files and the coordinate files (prmtop file and inpcrd file, respectively) were converted to the GROMACS topology files and the coordinate files (top file and gro file) using the ACPYPE program (30). Then, we performed the energy minimization for 5000 steps, followed by the heating process from 0 to 5 K during 10 ps under an NVT ensemble and from 5 to 298 K during 90 ps under an NPT ensemble ($P = 1$ bar, $T = 5, 50, 100, 150, 200, 225, 250, 275$ and 298 K, with each simulation time of 10 ps). The total simulation time of the heating process was 100 ps. Finally, the production run was performed under the NPT ensemble ($P = 1$ bar and $T = 298$ K) during 1 μs. In all MD simulations, the linear constraint solver (LINCS) algorithm (31) was adopted for constraints of the hydrogen atoms, and the Nosé–Hoover thermostat (32–34) and the Berendsen barostat (35) were adopted for the temperature and pressure regulations, respectively. The periodic boundary condition was applied by the particle mesh Ewald method (36), and the direct space cut-off distance was set to 10 Å. The van der Waals interactions were calculated using a switched cut-off between 8 and 10 Å. The simulation time step was set to 2 fs, and the coordinate data were recorded every 100 ps. All MD simulations were performed using the GROMACS version 5.0.4 program (Abraham, M.J., van der Spoel, D., Lindahl, E., Hess, B. and the GROMACS development team (2014) GROMACS User Manual version 5.0.4, www.gromacs.org). The last 500 ns MD trajectories were exclusively used for the following calculation.

Calculation of electronic coupling between FADH$^-$ and DEAC

To theoretically estimate the orientation factor (κ), 200 snapshots in the last 500 ns MD trajectories were randomly chosen, and the coordinate data of the isoallox-azine moiety of FADH$^-$ and DEAC in each snapshot were extracted, in a similar manner to the previous literature (37). The N10 of isoalloxazine was capped with a methyl group, while the linker and the amino acid residue moieties of DEAC were replaced with an H atom. The geometry optimization and the excited state calculation were performed using Gaussian16 (see Supplementary Figure S2 and Supplementary Tables S2 and S3), at the polarizable continuum model (PCM-) density functional theory (DFT) with B3LYP/6–311++G(2d,2p) and PCM- time-dependent (TD-) DFT with B3LYP/6–311++G(2d,2p) levels of theory, respectively, and the transition dipole moment magnitude of the donor (μ_D) and the acceptor (μ_A) vectors were calculated for each snapshot. The angles of the obtained μ_A and μ_D vectors were used for calculation of κ. The distance R_{DA} between the center of the mass of FADH$^-$ and DEAC was also extracted from each snapshot. In all calculations, PCMs for FADH$^-$ and DEAC were set up to the protein environment ($\varepsilon = 4.0$) and the water environment ($\varepsilon = 78.4$), respectively, where ε represents the dielectric constant.

Because the spectral overlap between FADH$^-$ and DEAC in the modified *Tth*CPDs should be common even though the modified amino acid residue is different, the difference in the FRET efficiency among the modified *Tth*CPDs is expectedly derived from the electronic coupling (V) in Supplementary Equation S1 (see Supplementary Computational details). The V value for each snapshot was calculated by substituting the obtained κ, μ_A, μ_D and R_{DA} values in Supplementary Equation S3.

RESULTS AND DISCUSSION

Design of protein-labeling oligonucleotides

We designed protein-labeling oligonucleotides, in which a UV lesion and a labeling nucleoside bearing a dye, through a reactive tosylate linker (25), were incorporated into the center of the oligonucleotides and the side part of the lesion, respectively, so the chromophore can be conjugated to an amino acid residue proximal to the active site of the PL, where FADH$^-$ is buried. Upon the formation of the DNA–protein complex, the active site is occupied by the lesion, whereas the flanking oligonucleotides with the labeling nucleoside can move around the active site. In the bound state, the chromophore on the labeling nucleoside is transferred from the DNA to a residue in the protein through an S$_N$2-type reaction (Figure 2A). The advantage of this labeling strategy is that (i) the recovery of the labeled proteins can be achieved by light illumination, because the affinity of DNA photolyases to non-damaged (namely, repaired) DNA is 7.5×10^4 times lower than that to UV-damaged DNA (38) and (ii) unlike conventional protein labeling using small molecules, undesirable labeling of the active site, which would lead to a reduction in the affinity of the protein to the damaged DNA, is avoidable. For the efficient FRET reaction, 7-diethylaminocoumarin (DEAC) was selected as the light-harvesting chromophore, as the emission spectrum of DEAC overlaps with the absorption spectrum of FADH$^-$ (Figure 2B). To search for the appropriate labeling positions proximal to FADH$^-$, we synthesized a series of oligonucleotides containing the *cis-syn* CPD as a
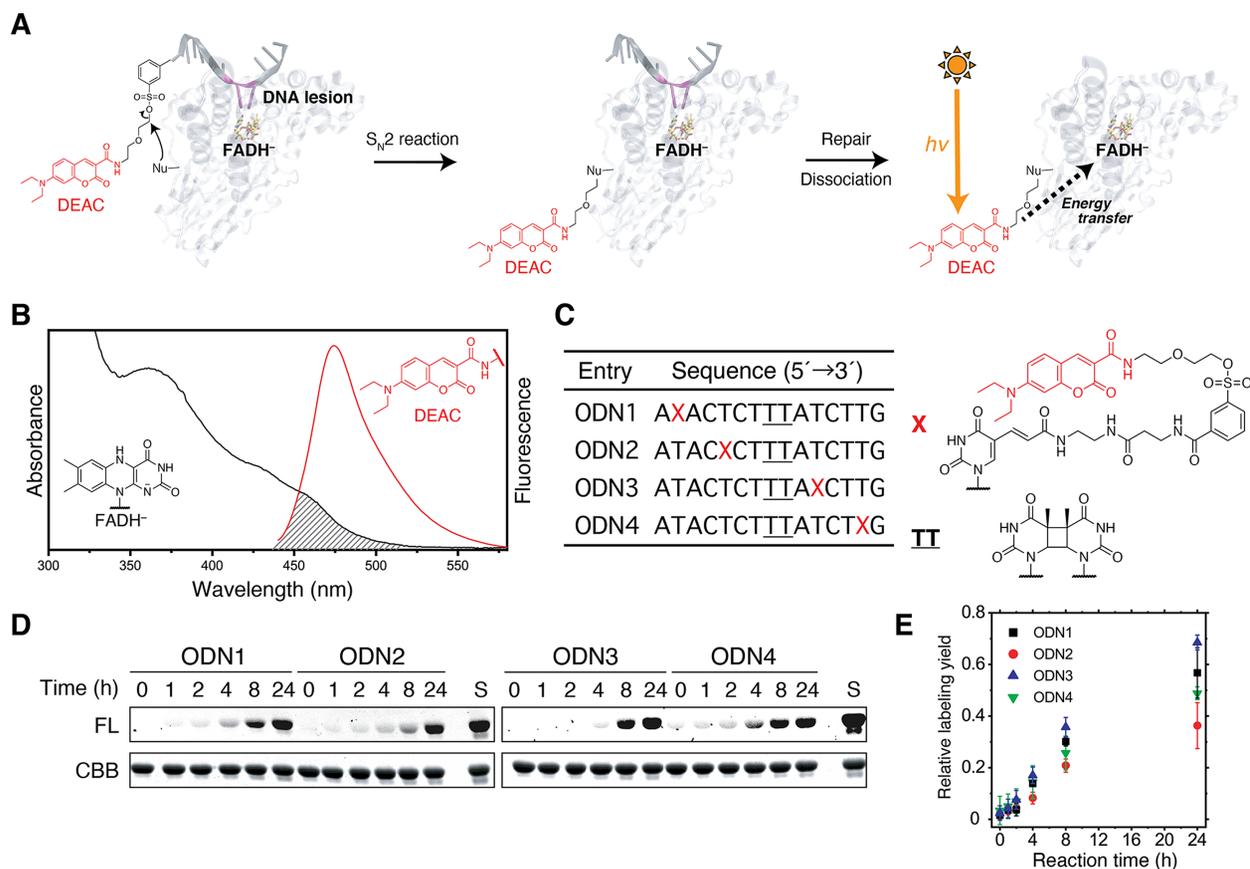
Figure 2. Labeling of DNA photolyase with 7-diethylaminocoumarin (DEAC). (**A**) Schematic illustration of the site-specific protein labeling strategy upon the complex formation of the protein-labeling oligonucleotides and *Tth*CPD, the structure of which was created with the reported *Tth*CPD structure (PDB ID: 2J09 (22)). The separation of the oligonucleotides from the protein is triggered by the DNA repair reaction. (**B**) Spectral overlap between the absorption spectrum of FADH⁻ and the fluorescence spectrum of DEAC. (**C**) Design of the site-specific protein-labeling oligonucleotides containing the modified nucleoside (X) bearing DEAC through the tosylate group and *cys-syn* CPD (TT). (**D**) In-gel fluorescence imaging (FL) and Coomassie Brilliant Blue (CBB) staining in the SDS-PAGE analysis of the labeling reaction. *Tth*CPD (20 μM) was incubated with each oligonucleotide (**ODN1**, **2**, **3** and **4**) (20 μM) in 100 mM phosphate buffer (pH 8) at 37°C, and aliquots of the reaction mixtures after 0–24 h reaction times were withdrawn and subjected to 10% SDS-PAGE. To calibrate the relative labeling yield internally, *Tth*CPD (20 μM) incubated with **ODN3** (40 μM) for 48 h (S) was also analyzed as a standard. (**E**) Relative labeling yields of **ODN1–4**. The relative fluorescence intensity was obtained from the in-gel fluorescence images shown in Figure 2D and plotted over the reaction times. The experiments were performed in triplicate (*n* = 3), and the mean values and standard deviations are shown.

ligand for CPD-PL, in which the labeling nucleoside containing DEAC through the tosylate group was incorporated at different positions in the sequence (**ODN1**, **2**, **3** and **4**) (Figure 2C), by using the NHS-DEAC molecule (Supplementary Figure S3). We chose *Tth*CPD as a target protein, because (i) its crystal structure with the natural antenna chromophore FMN has been reported (22), (ii) it is a thermostable enzyme allowing us to safely elevate the reaction temperature and (iii) the cysteine residue, which is a strong nucleophile and thus would potentially inhibit the site-selective reaction, is not present on the surface of the protein.

Preparation and characterization of coumarin-attached photolyase

To evaluate the labeling reactions using **ODN1–4**, *Tth*CPD was incubated with **ODN1–4** over a 0–24 h time course in the dark. The samples were subjected to SDS-PAGE, and the products were visualized by in-gel fluorescence imag-

ing and CBB staining (Figure 2D). As a result, all of these oligonucleotides were able to react with *Tth*CPD. Among them, **ODN3** showed the highest labeling yield after 24 h (69 ± 3%), whereas **ODN2** had the lowest efficiency (36±9%) (Figure 2E). The different labeling yields were ascribed to the differences in the relative distances between the tosylate linkers in **ODN1–4** and the reactive residues around the active site. The labeled *Tth*CPDs were recovered by white light illumination and purified on an affinity column to separate the labeling reagents from the protein (Figure 2A). To assess the enzymatic activities of the non-labeled *Tth*CPD (**PL**) and the *Tth*CPDs fluorescently-labeled by **ODN1**, **2**, **3** and **4** (**PL**ODN1, **PL**ODN2, **PL**ODN3 and **PL**ODN4), we performed the repair experiments with another single-stranded 14-mer substrate containing the *cis-syn* CPD. The mixture of *Tth*CPD and the substrate was illuminated with continuous light (430–800 nm) under anaerobic conditions, and the repair reaction was monitored by HPLC (Supplementary Figure S4). The numbers of repaired substrates per enzyme were evaluated and plotted against the irradi-
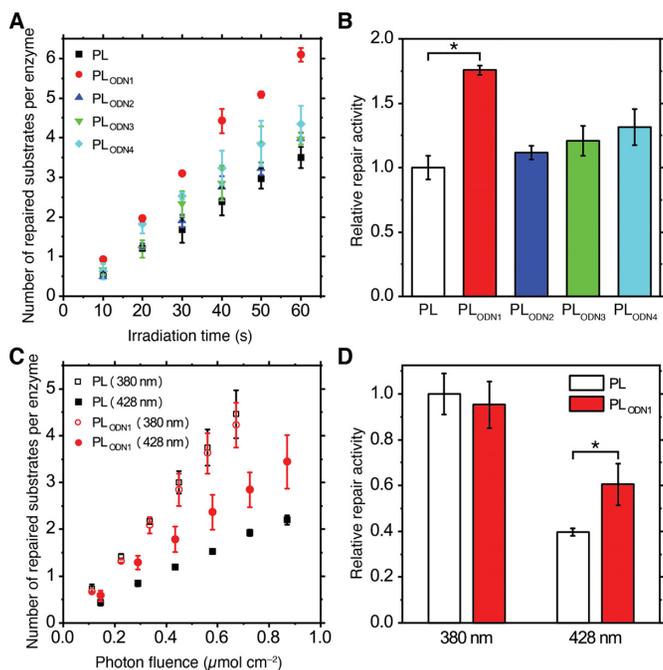
Figure 3. Repair activity of the fluorescently-labeled *Tth*CPDs. (**A**) The numbers of repaired substrates per enzyme for non-labeled *Tth*CPD (**PL**) and fluorescently-labeled *Tth*CPDs (**PL$_{ODN1}$**, **PL$_{ODN2}$**, **PL$_{ODN3}$** and **PL$_{ODN4}$**). The enzyme containing FADH$^-$ was prepared anaerobically and illuminated with white light (430–800 nm) in the presence of the single-stranded 14-mer substrate (5′-ATCGGCTTCGCGCA-3′). The progress of the repair reaction was monitored by HPLC. The experiments were performed in triplicate ($n = 3$), and the mean values and standard deviations are shown. (**B**) Relative repair activities of **PL** and the fluorescently-labeled *Tth*CPDs. The data shown in Figure 3A were fitted with a linear model, and the repair activity of **PL** was used as the standard to determine the relative repair activities of the fluorescently-labeled *Tth*CPDs. The statistical significance was evaluated by the two-sided Student's *t*-test, where the significance was set to $P < 0.05$. The asterisk indicates $P = 0.00040$. (**C**) The numbers of repaired substrates per **PL** and **PL$_{ODN1}$** plotted over the photon fluences (μmol cm^{-2}) upon illumination at either 380 nm or 428 nm. The experiments were performed in triplicate ($n = 3$), and the mean values and standard deviations are shown. (**D**) Relative repair activities of **PL** and **PL$_{ODN1}$** upon illumination at either 380 nm or 428 nm. The data shown in Figure 3C were fitted with a linear model, and the repair activity of **PL** at 380 nm was used as the standard to determine the relative repair activities of **PL** at 428 nm and **PL$_{ODN1}$** at 380 and 428 nm. The statistical significance was evaluated by the two-sided Student's *t*-test, where the significance was set to $P < 0.05$. The asterisk indicates $P = 0.0325$.

ation times (Figure 3A). The slopes of the linear regression were defined as the repair activities (Figure 3B). Consequently, all fluorescently-labeled *Tth*CPDs showed a higher repair activity than **PL**. Among the tested enzymes, **PL$_{ODN1}$** exhibited a 1.8-fold larger repair activity than **PL**, even though **PL$_{ODN3}$** was labeled with the chromophore in the highest yield (Figure 2E), implying that the fluorescently-labeled residue in **PL$_{ODN1}$** contributed remarkably to the efficient FRET from DEAC to FADH$^-$, as compared to that in **PL$_{ODN2-4}$**. It is, however, known that *Tth*CPD expressed in *E. coli* also loosely binds FMN at the secondary chromophore binding site (22), and therefore the enhancement of the repair activity of the fluorescently-labeled *Tth*CPDs may be ascribed to the presence of the natural secondary

Table 1. FMN contents of *Tth*CPD used for the repair experiment

Entry	FMN content
PL	28%
PL$_{ODN1}$	22%
PL$_{ODN2}$	18%
PL$_{ODN3}$	17%
PL$_{ODN4}$	30%

chromophore. To verify the possibility, the FMN content of the tested enzymes was analyzed by HPLC. FMN occupancies in the fluorescently-labeled *Tth*CPDs were found to be approximately 20–30%, at almost the same FMN amount of **PL** (Table 1). This result indicates that contribution of FMN to the enhanced DNA repair by **PL$_{ODN1-4}$** is almost identical and thus safely excludes the possibility that the enhanced repair activity of **PL$_{ODN1}$** was caused by the biased FMN occupation.

Additional repair experiments were performed under monochromatic light, at either 380 ± 5 nm or 428 ± 5 nm. If the energy transfer from DEAC to FADH$^-$ does occur, then **PL$_{ODN1}$** would be expected to repair the CPD more efficiently than **PL** upon illumination at 428 nm, which is close to the DEAC absorption maximum (Supplementary Figure S5). The obtained numbers of repaired substrates per enzyme with the monochromatic light sources were plotted over the incident photon fluences (Figure 3C and Supplementary Figure S6), and the repair activities of **PL** and **PL$_{ODN1}$** were evaluated (Figure 3D). As expected, **PL$_{ODN1}$** showed a 1.5-fold higher DNA repair activity with the 428 nm excitation, as compared to **PL**, whereas the repair activities of **PL** and **PL$_{ODN1}$** were comparable with the 380 nm excitation, where the extinction coefficients of FADH$^-$ and DEAC are similar. The results clearly demonstrated that the DNA repair activities of the fluorescently-labeled *Tth*CPDs were enhanced via the energy transfer from DEAC to FADH$^-$. Although the relative repair activity of **PL** at 428 nm was apparently reduced by 60% from that at 380 nm, this was caused simply by the difference in the extinction coefficients of FADH$^-$ at each wavelength (5600 and 2100 M^{-1} cm^{-1} at 380 and 428 nm, respectively), as the repair quantum yields of **PL** with the 380 and 428 nm excitations estimated by chemical actinometry (Supplementary Figures S6 and S7) were identical (54% and 57%, respectively).

To gain further insights into the origin of the enhanced repair activity via FRET, the positions of the labeled residues were determined by mass spectrometric peptide mapping, according to the procedure shown in Figure 4A. Peptides obtained after trypsin-digestion of the fluorescently-labeled *Tth*CPDs were analyzed by HPLC (Figure 4B), and the fractions with DEAC absorption were collected and analyzed by tandem mass spectrometry (Supplementary Figures S8–S23 and Table 2). In the case of **PL$_{ODN1}$**, the tandem mass analysis identified H123 as the major labeled residue, and H97, D128, and H257 as minor ones (Figure 4B). From the peak areas in the HPLC chromatograms, the relative labeling selectivities for the respective residues were found to be 6% for H97, 74% for H123, 8% for D128, and 12% for H257. The same procedure was applied to
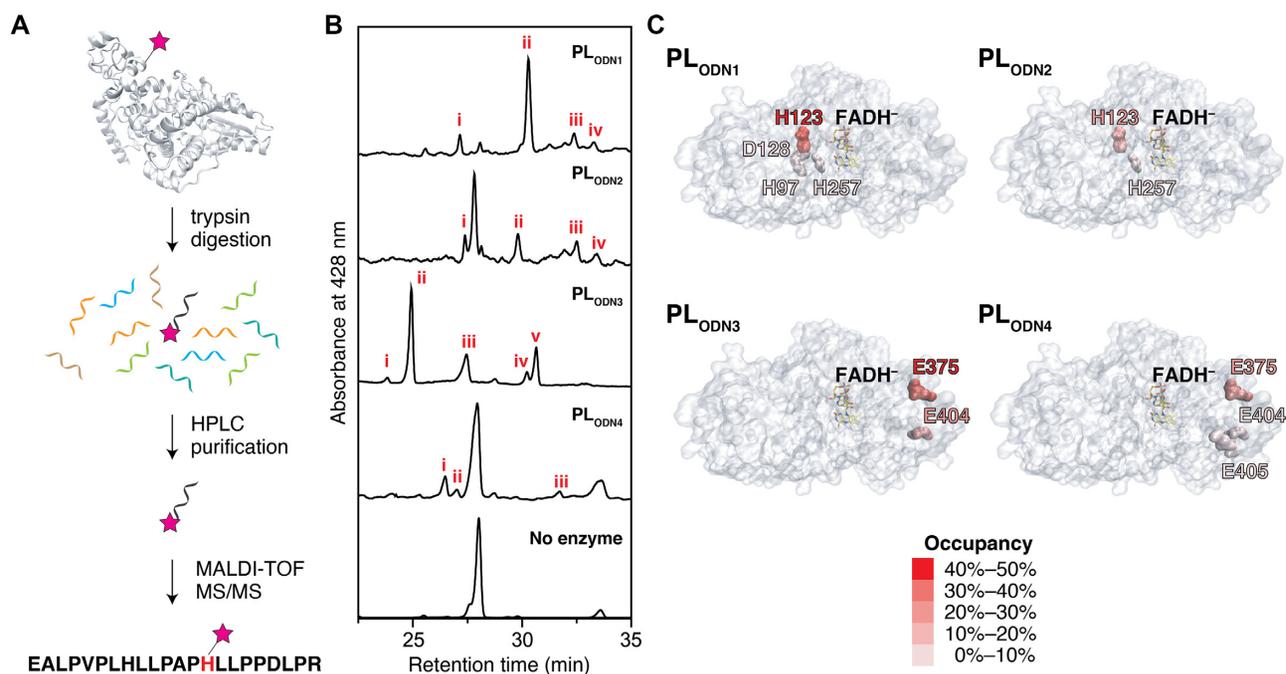
Figure 4. Peptide mapping of the fluorescently-labeled PLs. The structures in these figures came from the *Tth*CPD structure (PDB ID: 2J09 (22)). (**A**) Schematic view of the experimental procedure. (**B**) Analytical HPLC chromatograms of the digested products derived from fluorescently-labeled *Tth*CPDs. The presence of DEAC was monitored by absorption at 428 nm. The numbered peaks were collected and analyzed by MALDI TOF MS/MS. (**C**) Heat maps of the residues labeled with **ODN1–4**. DEAC occupancies for the respective amino acid residues were calculated by multiplying the labeling yield (Figure 2) and the residue selectivity (Table 1) and were visualized using heat maps.

PL$_{ODN2}$−PL$_{ODN4}$, and the relative labeling selectivities were also obtained (Table 2). In the MS/MS analysis of the fluorescently-labeled peptides, carbamylation of the N terminus of the peptide, the ε-amino group of lysine, and the guanidine moiety of arginine were observed. This was caused by the production of isocyanic acid during protein denaturation at high temperatures in the presence of urea (39–41), and therefore carbamylation of the labeled and recovered proteins was unlikely. As observed in the HPLC analysis of the sample with no enzyme, the products derived from the protein-labeling oligonucleotides were detected in the **PL$_{ODN2}$** and **PL$_{ODN4}$** samples (Figure 4B). This was consistent with the low labeling yields of **ODN2** and **ODN4** (Figure 2E) and indicated that the collision frequency of the tosylate and reactive amino acid side chain in the **PL-ODN2** and **PL-ODN4** complexes was much smaller than that in the **PL-ODN3** complex.

By taking into account the relative labeling yields and residue selectivities, the occupancies of the DEAC-labeled residues in **PL$_{ODN1}$−PL$_{ODN4}$** were evaluated and displayed using heat maps on the *Tth*CPD structure (Figure 4C). Obviously, the DEAC-labeled residues were clustered close together, depending on the position where the reactive tosylate linker was incorporated into the labeling oligonucleotides. This meant that our protein labeling strategy proceeded successfully, in a site-specific manner. As **PL$_{ODN1}$** exhibited the enhanced DNA repair activity, certain amino acid residues bearing DEAC in **PL$_{ODN1}$** would function as the light harvesting chromophore, and the involvement of the DEAC-attached residues in **PL$_{ODN2}$−PL$_{ODN4}$** in the FRET would be negligible.

Theoretical investigation of energy transfer from DEAC to FADH$^-$

According to the Förster theory, the FRET rate from donor to acceptor is expressed with i) the electronic coupling between donor and acceptor and ii) the spectral overlap integral between donor fluorescence and acceptor absorption (Supplementary Equation S1). In this study, the spectral overlaps between DEAC and FADH$^-$ were common in **PL$_{ODN1}$−PL$_{ODN4}$**, and therefore the electronic coupling between DEAC and FADH$^-$ would be the origin of the difference in the repair activities. To estimate the electronic coupling theoretically, we prepared model structures of *Tth*CPD bearing DEAC at the positions identified by the MS analysis, and performed 1 μs molecular dynamics (MD) simulations. As two imidazole nitrogens (N$^\delta$ and N$^\epsilon$) of histidine and two carboxy oxygens of aspartic acid can be conjugated with the DEAC moiety (termed type1 and type2, respectively, as shown in Figure 5A), the possible DEAC–amino acid conjugation patterns were considered (Supplementary Figure S1). After the MD simulations, 200 snapshots in the last 500 ns MD trajectories were randomly chosen and used for the electronic coupling calculations (Figure 5B). The distribution of the ingredients of the electronic coupling, namely amplitudes of transition dipole moments of donor and acceptor (μ_D and μ_A), distance between the donor and the acceptor (R_{DA}), and orientation factor κ^2 are summarized in Supplementary Figure S24. Among the tested systems, DEAC-attached H123 exhibited the highest electronic coupling (0.18 meV) regardless of the conjugation manner (Figure 5C), in good agreement with the

Table 2. Fluorescently-labeled peptides shown in Figure 4

Entry	Digested protein	Peak #	Sequence[a]	Labeled residue	Peak area
P1-1		i	DFSYHLLYHFPWMAERPLDPR	H257	12%
P1-2	PL$_{ODN1}$	ii	EALPVPLHLLPAPHLLPPDLPR	H123	74%
P1-3		iii	EALPVPLHLLPAPHLLPPDLPR[b]	D128	8%
P1-4		iv	AVYALTSHTPYGR	H97	6%
P2-1		i	DFSYHLLYHFPWMAERPLDPR	H257	19%
P2-2	PL$_{ODN2}$	ii	EALPVPLHLLPAPHLLPPDLPR	H123	52%
P2-2		iii	EALPVPLHLLPAPHLLPPDLPR[b]	H123	29%
P3-1		i	DPVVDLEEAR	E404	3%
P3-2		ii	VFNPVLQGER	E375	52%
P3-3	PL$_{ODN3}$	iii	WAPEYPSYAPKDPVVDLEEAR[b]	E404	22%
P3-4		iv	WAPEYPSYAPKDPVVDLEEAR[c]	E404	4%
P3-5		v	VFNPVLQGER[a]	E375	18%
P4-1		i	VFNPVLQGER	E375	59%
P4-2	PL$_{ODN4}$	ii	WAPEYPSYAPKDPVVDLEEAR[b]	E405	22%
P4-3		iii	WAPEYPSYAPKDPVVDLEEAR[c]	E404	19%
P4-4		iii	VFNPVLQGER[b]	E375	

[a]Red and green letters represent the fluorescently-labeled residues and the possible carbamylated residues, respectively.

[b]Carbamylated at either of the two positions in the sequence.

[c]Carbamylated at the two positions in the sequence.

0.2 meV value of the electronic coupling between FADH⁻ and the natural light-harvesting chromophore MTHF in the *E. coli* DNA photolyase structure (42). This would be attributable to the smaller R_{DA} and larger κ^2 than the others. In contrast, the DEAC-attached D128 exhibited the lowest electronic coupling among all conjugation systems regardless of the conjugation patterns, implying that DEAC-D128 would hardly contribute to the FRET. Indeed, the center value of κ^2 for DEAC-attached D128 was much smaller than the others. Unlike DEAC-attached H123 and D128, the estimated electronic couplings of DEAC-attached H97 and H257 dramatically varied depending on the conjugation patterns. This implied that these molecules would possibly contribute to the FRET depending on the relative configuration between FADH⁻ and DEAC. Given their low DEAC occupancy, their contribution would be limited. Considering the fact that H123 in *Tth*CPD was labeled by

ODN1 with the highest occupancy (Figure 4C), we concluded that the DEAC-labeled H123 is the main contributor to the enhancement of the DNA repair activity of PL$_{ODN1}$.

CONCLUSION AND PERSPECTIVES

We successfully enhanced the enzymatic activity of DNA photolyase by the introduction of an additional dye. As observed in the investigations of dye-sensitized solar cells, dye molecules have the potential to enhance the energy conversion efficiency (43). However, the direct involvement of an artificial dye in the conversion of light energy to chemical energy in biological phenomena has scarcely been reported to date. Thus, this work provides an example of the involvement of an artificial light-harvesting antenna chromophore in the enhancement of an enzymatic reaction. One may suppose that the 1.5-folds increase of the DNA re-
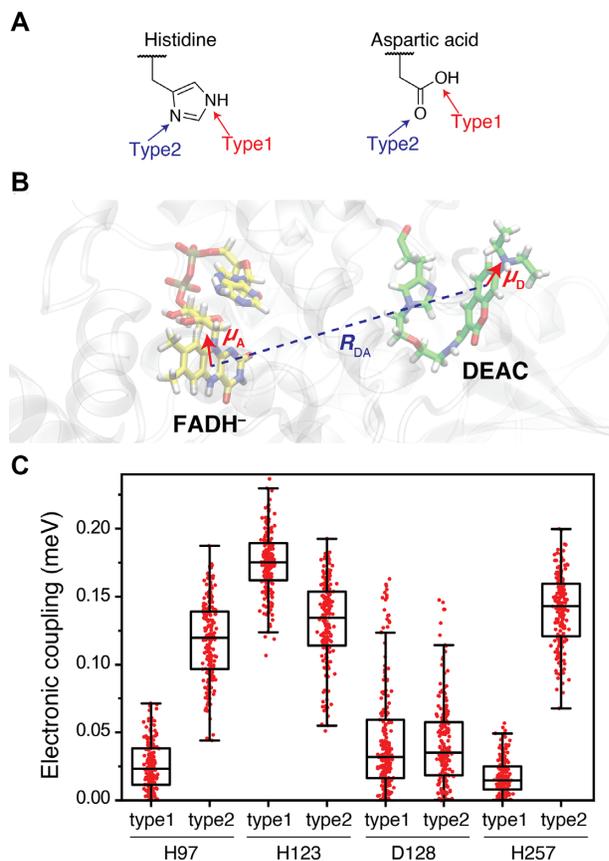
Figure 5. Calculated properties of the fluorescently-labeled residues. (**A**) Labeling patterns in the histidine and aspartic acid side chains. DEAC was covalently conjugated with either of the two imidazole nitrogens (N^δ and N^ε) in histidine and either of the two oxygen atoms in aspartic acid (termed type1 and type2). (**B**) Transition dipole moments of $FADH^-$ and DEAC illustrated in our model structure. A representative snapshot was extracted from the MD simulation of H123-type1. By using μ_D, μ_A and R_{DA} values shown in this panel, electronic coupling between DEAC and $FADH^-$ was calculated for each snapshot. (**C**) The electronic couplings between $FADH^-$ and DEAC, calculated using 200 snapshots in the last 500 ns MD trajectories. Each dot represents the result from each snapshot, and the overall results are shown in box-chart presentations, where the line in the boxes, the upper and lower limits of the boxes, and the whiskers represent the center value, the upper and lower quartiles, and the interquartile range, respectively. Note that the median values of the data set were similar to the center value.

pair activity caused by the coumarin modification is apparently weak. As observed in the tandem MS analysis, the coumarin-attached enzymes tested for the DNA repair were obtained as mixtures by the affinity-based labeling method (Table 1). We have suggested that H123 would be the main contributor of the enhanced DNA repair (Figure 5C), and therefore further improvement of the DNA repair activity of DNA photolyases will potentially be achieved by site-selective conjugation of various molecules at this position.

Since DNA binding proteins specifically bind their DNA substrates, the oligonucleotide-based protein modification shown in this study is widely applicable to labeling of proteins of interest. For instance, the pull-down assay using oligonucleotides is a powerful way to screen proteins that specifically bind to certain nucleic acids, such as UV-

damaged DNA, G-quadruplex, and i-motif DNA (44–46), but this strategy relies on high affinity of the proteins to these nucleic acids. Contrary to the canonical strategy, the reaction-based labeling of the protein would afford screening of proteins that exhibit relatively-weak affinity. Once the labeling reaction takes place and the proteins are tagged with the molecule during the association/dissociation process, the labeled proteins are traceable through the subsequent immunopurification step. Therefore, there are possibilities that unidentified nucleic acid-binding proteins that exhibit weak affinity could be found through this process.

Unlike conventional protein engineering, such as protein design and directed evolution (47,48), one of the advantages of the dye-assisted engineering of biological functions is that the expansion of the light absorbing property can be achieved by synthetic molecules. This method is expected to lead to the development of next-generation protein engineering and optogenetic tools by chemical approaches.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Sancar,A. (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem. Rev.*, **103**, 2203–2237.
2. Thiagarajan,V., Byrdin,M., Eker,A.P.M., Müller,P. and Brettel,K. (2011) Kinetics of cyclobutane thymine dimer splitting by DNA photolyase directly monitored in the UV. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 9402–9407.
3. Liu,Z., Tan,C., Guo,X., Kao,Y.-T., Li,J., Wang,L., Sancar,A. and Zhong,D. (2011) Dynamics and mechanism of cyclobutane pyrimidine dimer repair by DNA photolyase. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 14831–14836.
4. Yamamoto,J., Martin,R., Iwai,S., Plaza,P. and Brettel,K. (2013) Repair of the (6-4) photoproduct by DNA photolyase requires two photons. *Angew. Chem. Int. Ed.*, **52**, 7432–7436.
5. Thiagarajan,V., Villette,S., Espagne,A., Eker,A.P.M., Brettel,K. and Byrdin,M. (2010) DNA repair by photolyase: a novel substrate with low background absorption around 265 nm for transient absorption studies in the UV. *Biochemistry*, **49**, 297–303.
6. Payne,G. and Sancar,A. (1990) Absolute action spectrum of E-FADH₂ and E-FADH₂-MTHF forms of Escherichia coli DNA photolyase. *Biochemistry*, **29**, 7715–7727.

7. Lipman,R.S.A. and Jorns,M.S. (1992) Direct evidence for singlet-singlet energy transfer in Escherichia coli DNA photolyase. *Biochemistry*, **31**, 786–791.

8. Malhotra,K., Kim,S.-T., Walsh,C. and Sancar,A. (1992) FAD and 8-hydroxy-5-deazaflavin chromophores in photoreactivation by Anacystis nidulans DNA photolyase. *J. Biol. Chem.*, **267**, 15406–15411.

9. Schul,W., Jans,J., Rijksen,Y.M.A., Klemann,K.H.M., Eker,A.P.M., de Wit,J., Nikaido,O., Nakajima,S., Yasui,A., Hoeijmakers,J.H.J. *et al.* (2002) Enhanced repair of cyclobutane pyrimidine dimers and improved UV resistance in photolyase transgenic mice. *EMBO J.*, **21**, 4719–4729.

10. Johnson,J.L., Hamm-Alvarez,S., Payne,G., Sancar,G.B., Rajagopalan,K.V. and Sancar,A. (1988) Identification of the second chromophore of Escherichia coli and yeast DNA photolyases as 5,10-methenyltetrahydrofolate. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 2046–2050.

11. Scheerer,P., Zhang,F., Kalms,J., von Stetten,D., Krauß,N., Oberpichler,I. and Lamparter,T. (2015) The class III cyclobutane pyrimidine dimer photolyase structure reveals a new antenna chromophore binding site and alternative photoreduction pathways. *J. Biol. Chem.*, **290**, 11504–11514.

12. Yan,H., Zhu,K., Teng,M. and Li,X. (2020) A newly identified photolyase from Arthrospira platensis possesses a unique methenyltetrahydrofolate chromophore-binding pattern. *FEBS Lett.*, **594**, 740–750.

13. Ueda,T., Kato,A., Kuramitsu,S., Terasawa,H. and Shimada,I. (2005) Identification and characterization of a second chromophore of DNA photolyase from Thermus thermophilus HB27. *J. Biol. Chem.*, **280**, 36237–36243.

14. Fujihashi,M., Numoto,N., Kobayashi,Y., Mizushima,A., Tsujimura,M., Nakamura,A., Kawarabayasi,Y. and Miki,K. (2007) Crystal structure of archaeal photolyase from Sulfolobus tokodaii with two FAD molecules: Implication of a novel light-harvesting cofactor. *J. Mol. Biol.*, **365**, 903–910.

15. Kiontke,S., Gnau,P., Haselsberger,R., Batschauer,A. and Essen,L.-O. (2014) Structural and evolutionary aspects of antenna chromophore usage by class II photolyases. *J. Biol. Chem.*, **289**, 19659–19669.

16. Eker,A.P.M., Kooiman,P., Hessels,J.K.C. and Yasui,A. (1990) DNA photoreactivating enzyme from the cyanobacterium Anacystis nidulans. *J. Biol. Chem.*, **265**, 8009–8015.

17. Glas,A.F., Maul,M.J., Cryle,M., Barends,T.R.M., Schneider,S., Kaya,E., Schlichting,I. and Thomas,C. (2009) The archaeal cofactor F$_0$ is a light-harvesting antenna chromophore in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 11540–11545.

18. Zhang,F., Scheerer,P., Oberpichler,I., Lamparter,T. and 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. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7217–7222.

19. Kim,S.-T., Heelis,P.F., Okamura,T., Hirata,Y., Mataga,N. and 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**, 11262–11270.

20. Saxena,C., Sancar,A. and Zhong,D. (2004) Femtosecond dynamics of DNA photolyase: Energy transfer of antenna initiation and electron transfer of cofactor reduction. *J. Phys. Chem. B*, **108**, 18026–18033.

21. Tan,C., Guo,L., Ai,Y., Li,J., Wang,L., Sancar,A., Luo,Y. and Zhong,D. (2014) Direct determination of resonance energy pransfer in photolyase: Structural alignment for the functional state. *J. Phys. Chem. A*, **118**, 10522–10530.

22. Klar,T., Kaiser,G., Hennecke,U., Carell,T., Batschauer,A. and Essen,L.-O. (2006) Natural and non-natural antenna chromophores in the DNA photolyase from Thermus thermophilus. *ChemBioChem*, **7**, 1798–1806.

23. Morpurgo,M., Bayer,E.A. and Wilchek,M. (1999) N-hydroxysuccinimide carbonates and carbamates are useful reactive reagents for coupling ligands to lysines on proteins. *J. Biochem. Biophys. Methods*, **38**, 17–28.

24. Murata,T., Iwai,S. and Ohtsuka,E. (1990) Synthesis and characterization of a substrate for T4 endonuclease V containing a phosphorodithioate linkage at the thymine dimer site. *Nucleic Acids Res.*, **18**, 7279–7286.

25. Tsukiji,S., Miyagawa,M., Takaoka,Y., Tamura,T. and Hamachi,I. (2009) Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat. Chem. Biol.*, **5**, 341–343.

26. 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. *et al.* (2016) In: *Gaussian 09*, Revision D01, Gaussian, Inc., Wallingford.

27. Case,D.A., Betz,R.M., Cerutti,D.S., Cheatham,T., Darden,T., Duke,R.E., Giese,T.J., Gohlke,H., Goetz,A.W., Homeyer,N. *et al.* (2016) In: *AMBER 2016*. University of California, San Fransisco.

28. Maier,J.A., Martinez,C., Kasavajhala,K., Wickstrom,L., Hauser,K.E. and Simmerling,C. (2015) ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.*, **11**, 3696–3713.

29. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.

30. Vranken,W.F. and da Silva,A.W.S. (2012) ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes*, **5**, 367.

31. Hess,B., Bekker,H., Berendsen,H.J.C. and Fraaije,J.G.E.M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Compt. Chem.*, **18**, 1463–1472.

32. Nosé,S. (2006) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, **52**, 255–268.

33. Hoover,W.G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, **31**, 1695–1697.

34. Evans,D.J. and Holian,B.L. (1985) The Nose–Hoover thermostat. *J. Chem. Phys.*, **83**, 4069–4074.

35. Berendsen,H.J.C., Postma,J.P.M., van Gunsteren,W.F., DiNola,A. and Haak,J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684–3690.

36. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.

37. Kitoh-Nishioka,H., Yokogawa,D. and Irle,D. (2017) Förster resonance energy transfer between fluorescent proteins: Efficient transition charge-based study. *J. Phys. Chem. C*, **121**, 4220–4238.

38. Husain,I. and Sancar,A. (1987) Binding of E.coli DNA photolyase to a defined substrate containing a single T< >T dimer. *Nucleic Acids Res.*, **15**, 1109–1120.

39. Stark,G.R. (1965) Reactions of cyanate with functional groups of proteins. III. Reactions with amino and carboxyl groups. *Biochemistry*, **4**, 1030–1036.

40. Stark,G.R., Stein,W.H. and Moore,S. (1960) Reactions of the cyanate present in aqueous urea with amino acids and proteins. *J. Biol. Chem.*, **235**, 3177–3181.

41. Marier,J.R. and Rose,D. (1964) Determination of cyanate, and a study of its accumulation in aqueous solutions of urea. *Anal. Biochem.*, **7**, 304–314.

42. Park,H.-W., Kim,S.-T., Sancar,A. and Deisenhofer,J. (1995) Crystal structure of DNA photolyase from Escherichia coli. *Science*, **268**, 1866–1872.

43. Grätzel,M. (2003) Dye-sensitized solar cells. *J. Photochem. Photobiol. C: Photochem. Rev.*, **4**, 145–153.

44. Fujiwara,Y., Masutani,C., Mizukoshi,T., Kondo,J., Hanaoka,F. and Iwai,S. (1999) Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *J. Biol. Chem.*, **274**, 20027–20033.

45. Niu,K., Xiang,L., Jin,Y., Peng,Y., Wu,F., Tang,W., Zhang,X., Deng,H., Xiang,H., Li,S. *et al.* (2019) Identification of LARK as a novel and conserved G-quadruplex binding protein in invertebrates and vertebrates. *Nucleic Acids Res.*, **47**, 7306–7320.

46. Niu,K., Zhang,X., Deng,H., Wu,F., Ren,Y., Xiang,H., Zheng,S., Liu,L., Huang,L., Zeng,B. *et al.* (2018) BmILF and i-motif structure are involved in transcriptional regulation of BmPOUM2 in Bombyx mori. *Nucleic Acids Res.*, **46**, 1710–1723.

47. Kuhlman,B., Dantas,G., Ireton,G.C., Varani,G., Stoddard,B.L. and Baker,D. (2003) Design of a novel globular protein fold with atomic-level accuracy. *Science*, **302**, 1364–1368.

48. Arnold,F.H. (1998) Design by directed evolution. *Acc. Chem. Res.*, **31**, 125–131.