Crystal structure of cryptochrome 3 from *Arabidopsis thaliana* and its implications for photolyase activity

Yihua Huang*†, Richard Baxter*†, Barbara S. Smith*†, Carrie L. Partch†, Christopher L. Colbert†, and Johann Deisenhofer*†‡

*Howard Hughes Medical Institute and †Department of Biochemistry, University of Texas Southwestern Medical Center, 6001 Forest Park Road, Dallas, TX 75390

Contributed by Johann Deisenhofer, September 28, 2006 (sent for review September 8, 2006)

Cryptochromes use near-UV/blue light to regulate a variety of growth and adaptive process. Recent biochemical studies demonstrate that the Cryptochrome-*Drosophila, Arabidopsis, Synechocystis*, Human (Cry-DASH) subfamily of cryptochromes have photolyase activity exclusively for single-stranded cyclobutane pyrimidine dimer (CPD)-containing DNA substrate [Selby C, Sancar A (2006) *Proc Natl Acad Sci USA* 103:17696–17700]. The crystal structure of cryptochrome 3 from *Arabidopsis thaliana* (At-Cry3), a member of the Cry-DASH proteins, at 2.1 Å resolution, reveals that both the light-harvesting cofactor 5,10-methenyl-tetrahydrofolyl-polyglutamate (MTHF) and the catalytic cofactor flavin adenine dinucleotide (FAD) are noncovalently bound to the protein. The residues responsible for binding of MTHF in At-Cry3 are not conserved in *Escherichia coli* photolyase but are strongly conserved in the Cry-DASH subfamily of cryptochromes. The distance and orientation between MTHF and flavin adenine dinucleotide in At-Cry3 is similar to that of *E. coli* photolyase, in conjunction with the presence of electron transfer chain, suggesting the conservation of redox activity in At-Cry3. Two amino acid substitutions and the penetration of three charged side chains into the CPD-binding cavity in At-Cry3 alter the hydrophobic environment that is accommodating the hydrophobic sugar ring and thymine base moieties in class I CPD photolyases. These changes most likely make CPD binding less energetically favorable and, hence, insufficient to compete with pairing and stacking interactions between the CPD and the duplex DNA substrate. Thus, Cry-DASH subfamily proteins may be unable to stabilize CPD flipped out from the duplex DNA substrate but may be able to preserve the DNA repair activity toward single-stranded CPD-containing DNA substrate.

cryptochrome-DASH | DNA repair | light-harvesting cofactor

Photolyases/cryptochromes are blue-light photoreceptors, comprising a class of structurally related but functionally distinct flavoproteins found in all kingdoms of life (1–4). This family of proteins has been further divided into five subfamilies based on molecular phylogenetic analysis and function: (*i*) animal cryptochromes and (6-4) photolyases, (*ii*) class I cyclobutane pyrimidine dimer (CPD) photolyases, (*iii*) plant cryptochromes, (*iv*) class II CPD photolyases, and (*v*) the newly identified cryptochrome-*Drosophila*, *Arabidopsis*, *Synechocystis*, Human (Cry-DASH) cryptochromes (5).

Cryptochromes, which are similar in sequence and structure to photolyases but lack DNA repair activity, use near-UV/blue light to regulate a variety of growth and adaptive processes in organisms from bacteria to humans. For example, cryptochromes in *Arabidopsis thaliana* (termed At-Cry1 and At-Cry2) are nuclear proteins that mediate light control of stem elongation, leaf expansion, photoperiodic flowering, and the circadian clock by interacting with proteins such as phytochromes, COP1, and clock proteins, or/and chromatin and DNA (6, 7). In bacteria, *Drosophila*, and mammals, abundant evidence shows that cryptochromes are important components of the circadian-rhythm molecular machinery (3, 8–10). In *Synechocystis*, structural and functional studies suggest that Cry-DASH functions as a transcriptional repressor (5, 11).

DNA photolyases repair CPD lesions by using blue or near-UV light to catalyze the cleavage of the cyclobutane ring. The catalysis normally involves two noncovalently bound chromophores in the protein: the catalytic cofactor FAD and the light-harvesting cofactor, which is either 5,10-methenyl-tetrahydrofolyl-polyglutamate (MTHF) or 8-hydroxy-7,8-didemethyl-5-deazariboflavin (4, 12, 13). The crystal structures of three class I CPD photolyases from different microbes and the complex of *Anacystis nidulans* photolyase (termed An-photolyase) and a synthetic CPD DNA lesion have been determined (12–15). These structural studies, together with biochemical characterization, revealed important aspects of the structure and function of class I CPD photolyases, including the photocycle, the radical reaction mechanism, and substrate recognition (4).

Structures of the photolyase-like domain of Cry1 from *Arabidopsis thaliana* cryptochrome 1 (At-Cry1), a member of plant cryptochrome subfamily, and cryptochrome DASH from *Synechocystis* sp. PCC 6803 (Cry-DASH) have been reported (11, 16). Both structures contain FAD but no antenna pigment. These cryptochrome structures reveal general similarities such as FAD binding and the conservation of the overall domain architecture and arrangement, as well as structural differences such as changes in protein surface charge and the configuration of the CPD binding cavity between class I CPD photolyase and cryptochrome.

Recently, Kleine *et al.* (17) have identified the *Arabidopsis* gene At5g24850, termed At-Cry3, as a Cry-DASH member of the cryptochromes, closely related to Cry-DASH in Cyanobacterium *Synechocystis*. At-Cry3 carries an N-terminal sequence, which mediates import into chloroplasts and mitochondria. Biochemical characterization and spectroscopic analysis implies that, like *Escherichia coli* photolyase, At-Cry3 possesses the light-harvesting cofactor MTHF (18, 19). Biophysical studies by using femtosecond-resolved fluorescence upconversion techniques indicate that the energy-transfer process (60 ps) of cryptochrome 1 from *Vibrio cholerae* (VcCry1), a member of the Cry-DASH subfamily of cryptochromes, is four times faster than that in photolyase (292 ps) (20). In an accompanying manuscript, Selby and Sancar demonstrate that the Cry-DASH subfamily of

Author contributions: Y.H. designed research; Y.H. and B.S.S. performed research; C.L.P. contributed new reagents/analytic tools; Y.H. and C.L.C. analyzed data; and Y.H., R.B., and J.D. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviations: At-Cry, cryptochrome of *Arabidopsis thaliana*; CPD, cyclobutane pyrimidine dimer; Cry-DASH, cryptochrome-*Drosophila*, *Arabidopsis*, *Synechocystis*, Human; FAD, flavin adenine dinucleotide; MTHF, 5,10-methenyltetrahdrofolate.

Data deposition: X-ray crystallographic structure factors and coordinates for the refined model have been deposited to the Protein Data Bank, www.pdb.org (PDB ID code 2IJG).

‡To whom correspondence should be addressed. E-mail: johann.deisenhofer@utsouthwestern.edu.

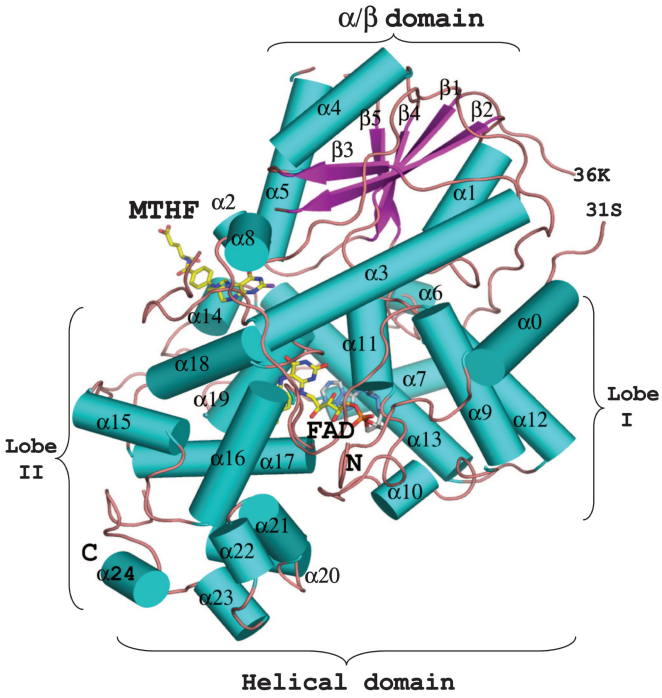

Fig. 1. Overall view of the structure of At-Cry3 in a schematic drawing: cyan, α helix; magenta, β-strands; orange, loop; MTHF and FAD are shown as stick models. The helical domain includes lobe I and lobe II.

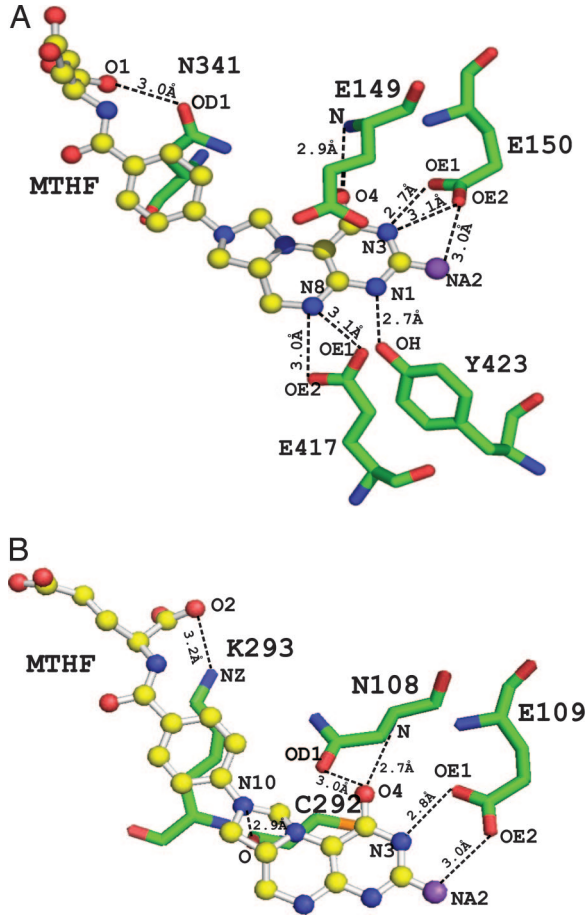

Fig. 2. Comparison of cofactor MTHF binding in At-Cry3 and *E. coli* photolyase. (*A*) Close-up view of MTHF binding in At-Cry3. (*B*) Close-up view of MTHF binding in *E. coli* photolyase. Possible H bonds are indicated in dashed lines, and distances are also labeled. MTHF are shown in ball and stick. Atom colors are as follows: yellow, carbons in MTHF; green, carbons in protein; red, oxygens; blue, nitrogens. Only residues that make hydrogen bonds to MTHF are shown.

cryptochromes actively repair single-stranded but not double-stranded CPD-containing DNA substrate (21).

To further elucidate general and subfamily specific aspects of photolyases/cryptochromes, we have determined the crystal structure of At-Cry3 at 2.1 Å resolution. Both MTHF and FAD are found to bind noncovalently to At-Cry3. The structure also reveals that alterations in the potential CPD-binding cavity of At-Cry3 may decrease the affinity to CPD. The polarity of the CPD-binding cavity is critical not only for the high affinity between protein and DNA substrate, but also crucial for the disruption of the Watson–Crick pairs formed between CPD and bases from complementary strand of the duplex DNA substrate. This observation may apply to other members of the Cry-DASH subfamily of cryptochromes. Thus, our results provide an explanation for the functional differences between class I CPD photolyases and Cry-DASH subfamily of cryptochromes on substrate recognition.

Results and Discussion

Overall Structure of At-Cry3. At-Cry3 crystals were obtained from 2.0 M ammonium sulfate, a crystallization condition very similar to that for Cry-DASH. We obtained two crystal forms ($P6_122$ and $P2_12_12_1$) from the same crystallization drop; both crystal forms had similar diffraction qualities (≈2.1-Å resolution). The following discussion is based on the structure determined from the data set of space group $P6_122$.

The positions of the secondary elements in the amino acid sequence are shown in Table 2, which is published as supporting information on the PNAS web site. The α/β domain of At-Cry3 is similar to that of other known structures of photolyase/cryptochrome proteins. The helical domain can be divided into two distinct lobes: Lobe I contains five long helices (α0, α9, and α11-α13), and a short helix α10. Lobe II consists of helices from α14 to α24 (Fig. 1).

In contrast to other cryptochrome/photolyase proteins of known structures, At-Cry3 contains an extra N-terminal fragment (residues 1–40). This fragment, containing one long α helix (α0), folds back to the cleft between lobe I and lobe II. Helix α0 makes extensive contacts with helices α9 and α12 of lobe I. The loop segments of this fragment also interact with the long loop that wraps the α/β domain and the helical domain of At-Cry3, further stabilizing its conformation (Fig. 1).

MTHF-Binding Site. MTHF-binding pocket is constructed from helices α5 and α14 and two loops (loop one connects helix α2 and α3; loop two connects helix α18 and α19). A total of 13 residues (His-84, Cys-147, Ser-148, Glu-149, Glu-150, Asn-341, Phe-344, His-345, Glu-417, Tyr-423, Pro-425, and Tyr-429) from these secondary structural elements interact with MTHF. MTHF makes eight hydrogen bonds to the side-chain atoms of Glu-149, Glu-150, Asn-341, Glu-417, and Tyr-423 (Fig. 2*A*). Among them, residues Glu-150 and Glu-417 contribute three and two hydrogen bonds, respectively. These hydrogen bonds, oriented in different directions, both enhance the affinity between MTHF and the protein and stabilize the conformation of MTHF.

Conservation of the MTHF-Binding Site in the Cry-DASH Subfamily. MTHF is not visible in the two available structures of cryptochromes, Cry-DASH, and At-Cry1 (14, 16). The sequences of At-Cry3 and Cry-DASH are 52% identical and the structures are

very similar, with an rms deviation of 1.14 Å for a structural superposition of 361 Cα atoms. Six of thirteen residues which have contacts with MTHF in At-Cry3 are identical in Cry-DASH. These residues, including Glu-149, Glu-150, Asn-341, Glu-417, and Tyr-423, contribute all of the hydrogen bonds for binding MTHF in At-Cry3 (the corresponding residues in Cry-DASH are as follows: Glu-113, Glu-114, Asn-305, Glu-381, and Tyr-387) (Fig. 6, which is published as supporting information on the PNAS web site). Moreover, the rmsd values of Cα atoms of each of these five pairs of residues are <0.6 Å. These observations suggest that Cry-DASH could bind MTHF. The absence of MTFH in the Cry-DASH crystal structure may result either from the loss of MTHF during purification or from insufficient supply of MTHF during the overexpression of recombinant Cry-DASH polypeptide.

Recent spectroscopic analysis indicates that cryptochrome from *V. cholerae* (Vc-Cry1) also binds MTHF (22). A sequence alignment between At-Cry3 and Vc-Cry1 suggests that MTHF could bind to both proteins in a similar way, because four of the five residues that contribute hydrogen bonds in At-Cry3/MTHF interaction are conserved (Asp-110, Glu-111, Glu-365, and Tyr-371 in Vc-Cry1). Alignment of all 10 sequences available for Cry-DASH subfamily cryptochromes indicates that four of the five residues that contribute hydrogen bonds to MTHF in At-Cry3 are identical in this subfamily of proteins (Fig. 6) (5). In contrast, none of these residues are conserved in At-Cry1, which is a member of the plant cryptochrome subfamily. We therefore hypothesize that Cry-DASH subfamily cryptochromes use MTHF as a light-harvesting cofactor and that the binding mode of MTHF is conserved in this subfamily.

The Binding Mode of MTHF Is Not Conserved Between At-Cry3 and *E. coli* Photolyase. *E. coli* photolyase and At-Cry3 are the only structures available of MTHF-binding proteins. The binding of MTHF to *E. coli* photolyase is different and appears much weaker compared with At-Cry3. In *E. coli* photolyase, only six residues (His-44, Glu-106, Asn-108, Glu-109, Cys-292, and Lys-293) make contacts to MTHF, and the contacts are largely polar interactions: Four residues contribute a total of six hydrogen bonds (Fig. 2*B*; ref. 13). The lower affinity of *E. coli* photolyase for MTHF explains the need to supply exogenous MTHF for crystallization of the holoenzyme (23). Most strikingly, none of the 13 residues in At-Cry3 that interact with MTHF is conserved in *E. coli* photolyase and vice versa. The MTHF-binding mode in At-Cry3 represents a previously undescribed cofactor–protein interaction, which is conserved in the Cry-DASH subfamily but distinct from that in *E. coli* photolyase.

FAD Binding and the Potential Electron Transfer Chain in At-Cry3. In At-Cry3, FAD adopts a U-shaped conformation with the isoalloxazine and adenine rings in close proximity to the solvent-accessible CPD cavity, similar to FADs in the other known structures of photolyase/cryptochrome proteins. The residues interacting with FAD in At-Cry3 are strongly conserved in all photolyase/cryptochrome proteins of known structures (11–13, 15, 16). In addition, an electron transfer chain from the protein surface to the FAD cofactor for activation of oxidized FADH· (Trp-306, Trp-359, and Trp-382 in *E. coli* photolyase) is also well conserved in At-Cry3 (the corresponding residues are Trp-356, Trp-409, and Trp-432, and the rmsd values of Cα atoms is <0.5 Å for all three pairs of residues) (11, 24, 25). The distance between the centers of mass of the atoms connected by conjugated double bonds of MTHF and of the isolloxazine ring observed in At-Cry3 structure is 15.2 Å, relatively shorter than that in *E. coli* photolyase (16.8 Å). Superposition of At-Cry3 structure with *E. coli* photolyase structure shows that the orientation between MTHF and FAD in two structures is very similar. These observations are consistent with biophysical measurements of the energy-transfer process in VcCry1 (20).

DNA Substrate Recognition of At-Cry3. To understand the functional differences between cryptochromes and class I CPD photolyase, we calculated the electrostatic surface potential for all six available structures of photolyase/cryptochrome proteins (26). Apart from At-Cry1, which has a predominantly negative electrostatic surface potential, all proteins possess a positive electrostatic surface potential around the potential CPD-binding cavity, implying a DNA-binding capability (ref. 17; Fig. 7, which is published as supporting information on the PNAS web site). This observation is in agreement with the experimental data: All five cryptochrome/photolyase proteins except At-Cry1 bind DNA in a sequence-independent manner. Moreover, it appears that both At-Cry3 and Cry-DASH have an even broader area of positive electrostatic surface potential than *E. coli* photolyase (S3. *A*, *B*, and *E*). Thus, the differences of DNA repair activity between Cry-DASH proteins and class I CPD photolyases are not related to DNA-binding capability.

To identify residues responsible for DNA substrate recognition in At-Cry3, we calculated all interactions between CPD-like DNA and An-photolyase. A total of 23 residues in An-photolyase were found to interact with CPD-like DNA. Among them, 13 residues interact with the CPD-containing strand, and 11 residues interact with the complementary strand of the DNA substrate, as observed by Mees *et al.* (14). Next, we superimposed At-Cry3 with the An-photolyase/CPD-like DNA complex structure to obtain new coordinates for At-Cry3. We then calculated the interaction between the new At-Cry3 coordinates and CPD-like DNA to identify residues that interact with DNA in At-Cry3. Residues interacting with CPD-like DNA in other photolyases/cryptochromes are identified in a similar way.

Residues That Interact with the CPD-Containing Strand Are Conserved Among Cry-DASH Proteins and Photolayses. We found that residues interacting with the CPD-containing strand are strongly conserved in the five photolyase/cryptochrome proteins known to bind DNA (except At-Cry1, which does not bind DNA) (Fig. 4). Of the 13 residues found to interact with the CPD-containing strand, six residues (Arg-232, Glu-283, Trp-286, Asn-349, Met-353, and Trp-392 in An-photolyase) directly interact with CPD (shown in red in Fig. 4). The other seven residues interact with the phosphate groups of the CPD-containing strand. Of these seven residues, the residues Gln-461 and Lys-414 may not be crucial for binding as replacements are found in Tt-photolyase. Mutations of residues (An-photolyase/Ec-photolyase: Gln-411/Gln-403, Lys-414/Lys-406) affected only the affinity toward CPD-containing DNA but not the quantum yield of CPD cleavage (27). Two other positions exhibit some variation: Gly-150 of An-photolyase is replaced by Thr in Ec-photolyase, Tt-photolyase, and At-Cry3. However, the interaction with the CPD-containing strand at this position occurs through the main chain N and is conserved. The other substitution of Asn-407 in An-photolyase for Ser-449 in At-Cry3 results in the loss of one hydrogen bond and is unlikely to affect the overall protein–DNA interaction. A sequence alignment of all 12 class I CPD photolyases and all 10 Cry-DASH subfamily cryptochromes extends the conservation of residues observed to bind the CPD-containing strand. In contrast, residues that interact with the complementary strand are not conserved between the three class I CPD photolyases, or in cryptochromes. However, the interactions between An-photolyase and the complementary DNA strand of the substrate primarily involve the main-chain atoms of the protein and the phosphate groups of the DNA. The binding mode of DNA substrate appears to be conserved between Cry-DASH subfamily proteins and class I CPD photolyases.

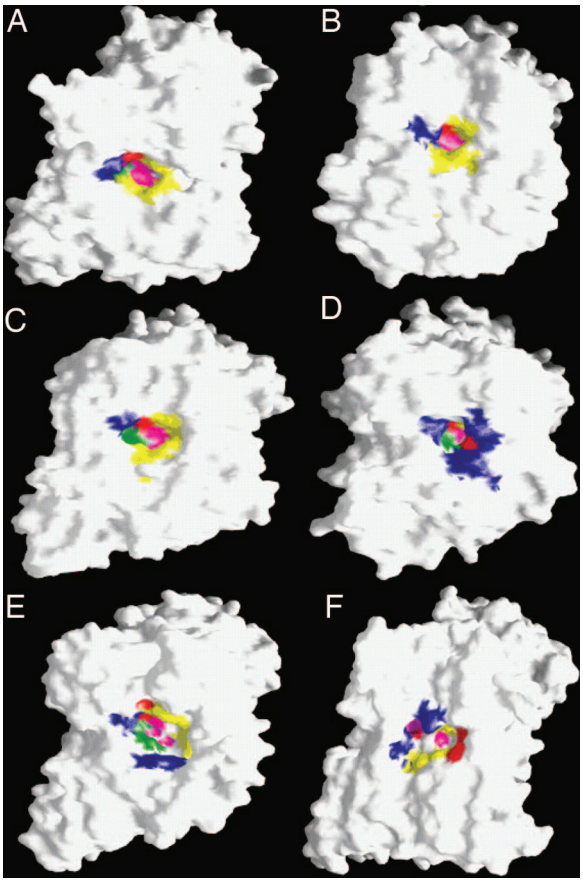

Fig. 3. Distribution of residues comprising the CPD binding cavities in photolyase/cryptochrome proteins. (*A*) An-photolyase (R232, E283, W286, N349, M353, W392, and FAD). (*B*) *E. coli* photolyase (R226, E274, W277, N341, M345, W384, and FAD). (*C*) Tt-photolyase (W247, W353, M314, R201, E244, N310, and FAD). (*D*) At-Cry3 (R443, W328, E444, R446, R276, N431, E325, Q395, N391, and FAD). (*E*) Cry-DASH (W292, R240, N395, N404, E289, R356, Q359, N355, and FAD). (*F*) At-Cry1 (Y402, L296, R360, V363, D409, S292, D359, and FAD). Residues are identified based on the calculation of a hypothetical complex between photolyase/cryptochrome and CPD (except for An-photolyase). The residues interacting with CPD are mapped on the surface of the protein with different colors: hydrophobic residues, yellow; negative residues, red; positive residues, blue; polar residues, green; FAD, magenta. All residues interacting with CPD in each of photolyase/cryptochrome proteins are mapped on the surface of the protein.

Residues Interacting with the CPD Are Strongly Conserved in Class I CPD Photolyase but Not in Cryptochromes. Six residues interact with the CPD: Arg-232, Glu-283, Trp-286, Asn-349, Met-353, and Trp-392 in the An-photolyase/CPD-like DNA complex (ref. 14; Fig. 4). The footprint of CPD, mapped on the surface of the protein, is shown in Fig. 3*A*. All six residues interacting with CPD in An-photolyase are identical in Tt-photolyase and *E. coli* photolyase. Furthermore, the rmsd values of Cα atoms of each pair are very small (<0.4 Å). In a sequence alignment of 12 known class I CPD photolyases from different species, residues Met-353, Trp-286, and Asn-349 are identical, and residues Trp-392, Arg-232, and Glu-283 are well conserved (data not shown) (5). Four of the six An-photolyase residues, Arg-232, Glu-283, Trp-286, and Asn-349, are conserved in At-Cry3 and the Cry-DASH subfamily. However, residues Met-353 and Trp-392 of An-photolyase are replaced by Gln-395 and Tyr-434, respectively. The same replacements occur in all 10 Cry-DASH subfamily proteins. In the 11 known plant cryptochrome subfamily members, residues Met-353 and Trp-392 are consistently

Fig. 4. Residues interacting with CPD-containing strand of DNA substrate in six photolyase/cryptochrome proteins. The fraction buried in the protein/DNA substrate complex of three important hydrophobic residues (M353, Trp-286, and Trp-392) configuring the CPD-binding cavity are calculated and shown in the figure. Residues marked in red interact directly with the CPD.

replaced by Val and Tyr, respectively. In the (6-4) photolyase subfamily, Met-353 is replaced consistently by His, but Trp-392 is identical (5). Because (6-4) photolyases recognize a substrate that has a different distribution of hydrophobic and hydrophilic groups, the replacement by a His residue may favor (6-4) photolyase recognition of its specific substrate and prevent the *cis*-syn CPD dimer from flipping into the active site (4). This analysis suggests a role for two hydrophobic residues (Met-353 and Trp-392 of An-photolyase) in differentiating class I CPD photolyases from the Cry-DASH subfamily of cryptochromes.

The Polarity of CPD-Binding Cavity Is Conserved in Class I CPD Photolyases but Not in Cryptochromes. Through calculation based on a hypothetical complex of cryptochrome/photolyase and a CPD-containing DNA substrate, we identified all residues likely to interact with the CPD in five other photolyase/cryptochrome proteins (residues are listed in Fig. 3). The distribution of CPD-contacting residues on the surface of these proteins, mapped with different colors, is displayed in Fig. 3 to show the polarity of the CPD-binding cavity. The distribution of the polarity within the CPD cavity is conserved in three class I CPD photolyases. Six residues (Trp-286, Met-353, Trp-392, Arg-232, Glu-283, and Asn-349 of An-photolyase and corresponding residues in other two class I CPD photolyases) form the majority of the concave surface of CPD-binding cavity. The position of these residues indicates the importance to the docking of polar CPD into the cavity. In three cryptochromes, extra residues contribute to the composition of the potential CPD-binding cavity surface, and the cavity polarity is disturbed. For example, in At-Cry3, a predominantly charged cavity surface replaces the hydrophobic surface of photolyases. This alteration of cavity polarity may decrease significantly the binding energy between CPD and CPD-binding cavity of At-Cry3 (Fig. 3).

Residue Substitution Within the CPD-Binding Cavity Suggest a Lower Affinity for the Cry-DASH Subfamily of Cryptochromes. In An-photolyase, Trp-286, Met-353, and Trp-392 form a hydrophobic region that makes extensive contact to the deoxyribose sugar and the rear edge of the pyrimidine rings of the CPD (Fig. 5). Residues Met-353 and Trp-286 form hydrophobic surfaces on either of the two pyrimidine rings that align with their planes parallel to each other and to Trp-286. At the front of the cavity, the side chains of three polar residues Glu-283, Arg-232, and Asn-349 make hydrogen bonds and polar interactions to the O2, N3, and O4 atoms of the pyrimidine rings to stabilize the conformation of the thymine dimer. The FAD sits beneath the three polar residues, with adenine atom AN6 in close proximity to the two O4 atoms of the CPD. At the back of the CPD cavity, Trp-392 forms a hydrophobic surface, making extensive van der

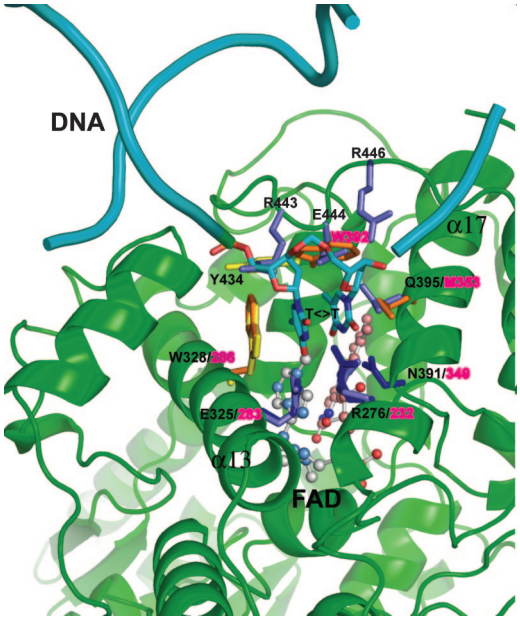

Fig. 5. Comparison of the residues interacting with CPD in An-photolyase and At-Cry3. Colors: At-Cry3, green; hydrophobic residues, yellow (W328 and Y434); charged/polar residues, light blue (R276, E325, N391, Q395, R443, E444, and R446); DNA backbone, cyan; CPD shown in stick with atom color: carbon, cyan; oxygen, red; nitrogen, blue. FAD shown in ball and stick style. Residues interacting with CPD in At-photolyase are colored: hydrophobic residues, brown (W286, W392, and M353); charged/polar residues, blue (E283, R232, and N349). The positions of these residues based on superposition An-photolyase with At-Cry3. Residues labeled in red are residues of An-photolyase (for clarity, An-photolyase structure was omitted).

Table 1. Crystallographic and refinement statistics

Measurement	Value
X-ray data	
Space group	$P6_1\ 22$
a and *b*, Å	$a = b = 117.897$
c, Å	$c = 216.567$
α, β, γ, °	$\alpha = \beta = 90, \gamma = 120$
Resolution, Å	50-2.1
No. of reflections (total)	52,633
Unique of reflections	52,553
No. of reflections used	50,639
Completeness,* %	100 (99.9)
Avg. I/σ	14.64 (2.13)
R_{sym}†	0.193 (0.873)
Refinement	
Resolution, Å	20-2.1
No. of protein atoms	4,069
No. of water atoms	231
R_{work},‡ %	23.3
R_{free},§ %	25.6
Bond lengths, Å	0.0063
Bond angle, °	1.49
Favored, %	95.6
Allowed, %	100

*Numbers in parentheses are for the highest-resolution shell of the data.
†$R_{sym} = \Sigma_h \Sigma_i |\bar{I}(h) - I_i(h)| / \Sigma_h \Sigma_i I_i(h)$, where $I_i(h)$ is the *i*th measurement of the intensity for Miller indices *h*, and $I(h)$ represents the mean intensity value of the symmetry (or Friedel) equivalent reflections of Miller indices *h*.
‡$R_{work} = \Sigma_h ||F_o| - |F_o|| / \Sigma_h |F_o|$.
§The formula for R_{free} is the same as that for R_{work}, except that it is calculated with 5% of the structure factors that had not been used for refinement.

Waals contact with the methyl group of one of the thymines (Fig. 5). Furthermore, Met-353 and Trp-286 of An-photolyase are completely buried in the An-photolyase/CPD-like DNA complex (Fig. 4). We speculate that these two residues must be strictly conserved in class I CPD photolyases and that mutation of either of these residues will significantly decrease DNA repair activity on duplex DNA substrate. On the other hand, Trp-392 is partially buried and only interacts with the methyl group of one thymine, allowing for limited conservative replacement in different class I CPD photolyases.

In AtCry3, the side chain of Tyr-434 in At-Cry3 has moved away from the position that Trp-392 adopts in An-photolyase, and the side chain of Glu-444 takes its place. Two other charged residues Arg-443 and Arg-446 become exposed to the CPD-binding cavity in a region that would otherwise be hydrophobic (Fig. 5). Also, the polar residue Gln-395 of At-Cry3 replaces Met-353 of An-photolyase within the CPD-binding cavity. These alterations to the CPD-binding cavity convert a hydrophobic environment to a more polar one, which may be less favorable for binding of the CPD in the active site.

Similarities and Differences of Cry-DASH and Class I CPD Photolyases. Class I CPD photolyases fulfill three requirements to perform their function of repairing CPD lesions: (*i*) they bind their substrate with a high affinity ($K_M \approx 10^{-9}\ M^{-1}s^{-1}$), (*ii*) they carry out the desired reaction rapidly and with high efficiency due to the specific position and environment of the bound cofactors, and (*iii*) they have significantly lower affinity for the product (unmodified DNA) than the reactant.

In an accompanying manuscript, Selby and Sancar demonstrate that Cry-DASH proteins actively repair single-stranded but not double-stranded DNA (21). They show that, like class I CPD photolyases, Cry-DASH proteins bind damaged ssDNA with a high affinity, and that they exhibit significantly lower affinity for the product unmodified ssDNA. We show here that, like *E. coli* photolyase, At-Cry3 binds MTHF and that the orientation and distance between it and the bound FAD are similar. This observation, and the conservation of the FAD environment and those residues of the active site with proposed catalytic activity (An-Photolyase: Glu-283 and Trp-286), strongly suggests that Cry-DASH proteins should have a similar catalytic rate and quantum yield for repair of bound CPD lesions. Hence, with respect to ssDNA, Cry-DASH proteins share all three requirements of enzymatic activity with class I CPD photolyases.

The essential difference between Cry-DASH proteins and class I CPD photolyases is the inability of the former to repair dsDNA. The reason behind this difference must lie in the different ability to efficiently bind the substrate, specifically the inability of Cry-DASH proteins to stabilize CPD lesions "flipped out" from the duplex DNA. Our analysis of all available sequences and structures show that Cry-DASH proteins share the residues and tertiary structure for binding to the phosphate backbone of the CPD-containing strand but that the mutation of two critical residues in the active site and an overall decrease in hydrophobicity are likely to lower the affinity of Cry-DASH proteins for the CPD itself. In the case of ssDNA, this lower affinity is of no consequence as the substrate does not have an alternative binding partner, but in the case of dsDNA, we propose that the free energy of association for Cry-DASH and the CPD lesion is less favorable than the alternative stacking and pairing interactions of the CPD within the DNA duplex (28), hence, the protein is unable to effectively bind its substrate. Further structural, functional, and computational studies shall elucidate which, if any, of these features may contribute to the differing substrate specificities between class I CPD photolyases and the Cry-DASH family of proteins.

BIOCHEMISTRY

Materials and Methods

UV-Visible (UV-Vis) Spectroscopy. For detailed information regarding the expression and purification of At-Cry3, please refer to *Supporting Text*, which is published as supporting information on the PNAS web site. To determine the presence of cofactors, UV-Vis absorption spectra of purified At-Cry3 were recorded by using a UV-2401 PC spectrophotometer (Shimadzu, Kyoto, Japan). The absorption spectrum of purified protein was dominated by a 384-nm peak, suggesting the presence of MTHF. Shoulders at 450 and 480 nm are typical for an oxidized FAD species, and minor peaks at 590 and 640 nm indicate the presence of low percentage of the neutral flavin radical FADH (Fig. 8, which is published as supporting information on the PNAS web site) (4, 18, 19).

Crystallization, Data Collection, and Structure Determination. Preliminary crystallization conditions were screened by the hanging-drop vapor-diffusion method. Crystals were observed in 2.0 M $(NH_4)_2SO_4$ (Hampton Screen I no. 32), and grew in 1–2 weeks, usually as single columns with green-yellowish color. Only fresh protein could be crystallized in 2.0 M ammonium sulfate. A complete data set to 2.1 Å was collected from a single frozen crystal at beamline 19-ID at the Structural Biology Center of the Advanced Photon Source at the Argonne National Laboratory (Argonne, IL). The data were indexed, integrated, and scaled by using HKL2000 package (29). The structure was determined by using molecular replacement protocols in Phaser (30). Atomic coordinates of Cry-DASH, PDB ID code 1NP7 without water or cofactor, were used as the molecular replacement model. A single monomer of At-Cry3 was located in one asymmetric unit. The model was built and adjusted in O (31) and refined by using rigid body refinement, simulated annealing, conjugated gradient minimization, and individual B factor minimization in CNS (32). Statistics for data collection and refinement are presented in Table 1.

We thank Dr. Aziz Sancar for At-Cry3 cDNA and Drs. Aziz Sancar and Chad Brautigam for valuable suggestions. The x-ray diffraction data were collected at the Argonne National Laboratory's Structural Biology Center at the Advanced Photon Source, which is supported by the Office of Energy Research, U.S. Department of Energy. This research is supported in part by Welch Foundation Grant I-1185 (to J.D.). J.D. is an Investigator with the Howard Hughes Medical Institute.

1. Cashmore AR, Jarillo JA, Wu YJ, Liu D (1999) *Science* 284:760–765.
2. Todo T (1999) *Mutat Res* 434:89–97.
3. Lin C, Shalitin D (2003) *Annu Rev Plant Biol* 54:469–496.
4. Sancar A (2003) *Chem Rev* 103:2203–2237.
5. Daiyasu H, Ishikawa T, Kuma K, Iwai S, Todo T, Toh H (2004) *Genes Cells* 9:479–495.
6. Yang HQ, Tang RH, Cashmore AR (2001) *Plant Cell* 13:2573–2587.
7. Somers DE, Devlin PF, Kay SA (1998) *Science* 282:1488–1490.
8. Ceriani MF, Darlington TK, Staknis D, Mas P, Petti AA, Weitz CJ, Kay SA (1999) *Science* 285:553–556.
9. Chen M, Chory J, Fankhauser C (2004) *Annu Rev Genet* 38:87–117.
10. Panda S, Hogenesch JB, Kay SA (2002) *Nature* 417:329–335.
11. Brudler R, Hitomi K, Daiyasu H, Toh H, Kucho K, Ishiura M, Kanehisa M, Roberts VA, Todo T, Tainer JA, *et al.* (2003) *Mol Cell* 11:59–67.
12. Tamada T, Kitadokoro K, Higuchi Y, Inaka K, Yasui A, de Ruiter PE, Eker AP, Miki K (1997) *Nat Struct Biol* 4:887–891.
13. Park HW, Kim ST, Sancar A, Deisenhofer J (1995) *Science* 268:1866–1872.
14. Mees A, Klar T, Gnau P, Hennecke U, Eker AP, Carell T, Essen LO (2004) *Science* 306:1789–1793.
15. Komori H, Masui R, Kuramitsu S, Yokoyama S, Shibata T, Inoue Y, Miki K (2001) *Proc Natl Acad Sci USA* 98:13560–13565.
16. Brautigam CA, Smith BS, Ma Z, Palnitkar M, Tomchick DR, Machius M, Deisenhofer J (2004) *Proc Natl Acad Sci USA* 101:12142–12147.
17. Kleine T, Lockhart P, Batschauer A (2003) *Plant J* 35:93–103.
18. Pokorny R, Klar T, Essen LO, Batschauer A (2005) *Acta Crystallogr F* 61:935–938.
19. Song SH, Dick B, Penzkofer A, Pokorny R, Batschauer A, Essen LO (2006) *J Photochem Photobiol B* 85:1–16.
20. Saxena C, Wang H, Kavakli IH, Sancar A, Zhong D (2005) *J Am Chem Soc* 127:7984–7985.
21. Selby C, Sancar A (2006) *Proc Natl Acad Sci USA* 103:17696–17700.
22. Worthington EN, Kavakli IH, Berrocal-Tito G, Bondo BE, Sancar A (2003) *J Biol Chem* 278:39143–39154.
23. Park HW, Sancar A, Deisenhofer J (1993) *J Mol Biol* 231:1122–1125.
24. Li YF, Heelis PF, Sancar A (1991) *Biochemistry* 30:6322–6329.
25. Aubert C, Mathis P, Eker AP, Brettel K (1999) *Proc Natl Acad Sci USA* 96:5423–5427.
26. Petrey D, Honig B (2003) *Methods Enzymol* 374:492–509.
27. Vande Berg BJ, Sancar GB (1998) *J Biol Chem* 273:20276–20284.
28. Park H, Zhang K, Ren Y, Nadji S, Sinha N, Taylor JS, Kang C (2002) *Proc Natl Acad Sci USA* 99:15965–15970.
29. Otwinowski Z, Minor W (1997) *Methods Enzymol* 276:307–326.
30. Bailey S (1994) *Acta Crystallogr D* 50:760–763.
31. Jones TA, Zou JY, Cowan, SW, Kjeldgaard (1991) *Acta Crystallogr A* 47:110–119.
32. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, *et al.* (1998) *Acta Crystallogr D* 54:905–921.