Evidence for Dinucleotide Flipping by DNA Photolyase*

Brian J. Vande Berg§ and Gwendolyn B. Sancar§

From the Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599-7260

DNA photolyases repair pyrimidine dimers via a reaction in which light energy drives electron donation from a catalytic chromophore, FADH²⁻, to the dimer. The crystal structure of Escherichia coli photolyase suggested that the pyrimidine dimer is flipped out of the DNA helix and into a cavity that leads from the surface of the enzyme to FADH²⁻. We have tested this model using the Saccharomyces cerevisiae Phr1 photolyase which is >50% identical to E. coli photolyase over the region comprising the DNA binding domain. By using the bacterial photolyase as a starting point, we modeled the region encompassing amino acids 383–530 of the yeast enzyme. The model retained the cavity leading to FADH²⁻ as well as the band of positive electrostatic potential which defines the DNA binding surface. We found that alanine substitutions at sites within the cavity reduced both substrate binding and discrimination, providing direct support for the dinucleotide flip model. The roles of three residues predicted to interact with DNA flanking the dimer were also tested. Arg⁴⁵² was found to be particularly critical to substrate binding, discrimination, and photolysis, suggesting a role in establishing or maintaining the dimer in the flipped state. A structural model for photolyase-dimer interaction is presented.

Pyrimidine bases absorb strongly in the UV region and are highly susceptible to photochemical reactions that alter their structures. In DNA, cyclobutane pyrimidine dimers (CPDs)¹ and pyrimidine-pyrimidone (6-4) photoproducts are the most frequent and biologically significant products of these reactions. These lesions are lethal and mutagenic and must be repaired to ensure cell survival and genetic stability. DNA photolyases repair CPDs and (6-4) photoproducts via reactions in which near UV or visible light provides the energy for bond breaking, resulting in a quantum yield for the overall photolysis reaction of 0.6–1.0 for CPDs containing thymine or uracil (6, 7).²

CPD photolyases are structure-specific enzymes that display binding discrimination comparable to that seen for sequence-specific DNA-binding proteins. Studies on the Escherichia coli and Saccharomyces cerevisiae enzymes have shown that the equilibrium association constant for (cis,syn)-CPDs in DNA is approximately 10⁹ M⁻¹, whereas the association constant for nondamaged DNA is 10⁷ M⁻¹ (6, 8–10). The affinity for (trans,syn)-pyrimidine dimers in DNA and U<−>U dimers in RNA is only about 10-fold greater than that for nondamaged DNA; nevertheless, once bound, these lesions are photolyzed efficiently (7, 11). Thus the presence of a cyclobutane dimer, the geometry of the bases in the dimer, and the absence of a 2′-OH on the sugar phosphate backbone are determinants of binding specificity. Important recognition elements are also found in DNA flanking the dimer. In particular, ethylation of the first phosphate 5′ to the dimer and 3–4 phosphates 3′ to the dimer in the lesion-containing strand inhibits binding (12, 13). At least some of these interactions contribute to binding specificity as shown by the fact that discrimination between dimer-containing and undamaged oligonucleotides decreases as the substrate is shortened (6). In addition, mutations in the yeast Phr1 photolyase have been identified which simultaneously decrease substrate discrimination and alter interactions with DNA phosphates surrounding the dimer (9). These results imply that the structure of the flanking sugar-phosphate backbone is uniquely altered by the dimer.

A structural basis for the efficiency of the photolysis reaction and for specific substrate recognition has been provided by the crystal structure of E. coli photolyase (14). The polypeptide chain is folded into an amino-terminal αβ domain and a carboxyl-terminal helical domain with the folate cofactor nestled into a shallow cleft between the two domains. The flavin chromophore lies deeply buried in the center of the helical domain, dipyrimidine photolyases, hereafter referred to as CPD photolyases, are the subject of this report.

Understanding how the CPD photolyases efficiently repair pyrimidine dimers in DNA entails answering the following two questions: how do the enzymes recognize pyrimidine dimers specifically in the midst of a vast excess of nondamaged bases, and how do the enzymes catalyze dimer photolysis? Photolysis involves two noncovalently bound chromophores, reduced FAD and a second chromophore which, depending upon the source of the enzyme, is either folate or deazaflavin (4). Absorbance of a photon of photoreactivating light subsequent to substrate binding initiates electron donation by enzyme-bound FADH²⁻ to one of the bases in the dimer (4, 5). The photon may be absorbed either directly by FADH²⁻ or, more often, by the second chromophore which transfers energy to the flavin chromophore (1). Both electron transfer and energy transfer are highly efficient processes, resulting in a quantum yield for the overall photolysis reaction of 0.6–1.0 for CPDs containing thymine or uracil (6, 7).²

* This work was supported by Grant GM35123 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed. Tel.: 919-966-2077; Fax: 919-966-2302; E-mail: gsanca.biochem@mhs.unc.edu.

1. The abbreviations used are: CPDs, cyclobutane pyrimidine dimers; FAD, flavin adenine dinucleotide; Phr1, photolyase encoded by the S. cerevisiae PHR1 gene; EMSA, electrophoretic mobility shift assay; LPCR, polymerase chain reaction; bp, base pair(s).

2. A. Sancar, personal communication.

This is an Open Access article under the CC BY license.
carbon of each amino acid were placed in the position of the previous E. coli amino acid. The yeast amino acid side chain was then rotated about its first dihedral rotation angle ($\chi_1$) to produce an alignment devoid of steric clash. Because $\chi_1$ displays a strong tendency to assume values near 60, 180, and 300$^\circ$, each new side chain was examined first in each of these positions. The preferred alignment was at a $\chi_1$ value that placed the yeast side chain near the E. coli side chain and did not produce any steric clashes (within 2.5 Å). At positions 285, 291, and 368 (Phr1 positions 395, 401, and 478), alternative values for $\chi_1$ had to be chosen. A similar protocol was used to model the dihedral angle $\chi_2$ between the $\beta$ and $\gamma$ carbons on Glu, Lys, Arg, Met, Phe, Tyr, and Thr side chains. Again, the preferred dihedral values were those that placed the yeast side chain near the corresponding E. coli side chain. Steric clashes at positions 307, 308, and 411 (Phr1 positions 417, 418, and 521) prohibited placing these side chains in the preferred positions. The program REFINEx, written by Dr. Jan Hermans (University of North Carolina, Chapel Hill), was then used to refine the geometry (bond lengths, bond angles, and dihedral angles) for the S. cerevisiae model. Individual amino acids were first refined, then the entire structure was analyzed over the course of 10 cycles. After 10 cycles, the cumulative coordinate shift was determined, and further modeling was performed until this shift value decreased to less than 5.0 Å. An identical approach was used to replace Arg$^{512}$, which lies near the rim of the cavity leading to flavin, with the equivalent yeast residue, Lys$^{330}$. The entire PHR1 coding region and approximately 500 bp of 3'-flanking sequence were subcloned from pCB12144/recon (9) into XmaI-Pst1-digested pmal-c2 (New England Biolabs), yielding pGBS424. Oligonucleotides used in the construction of the various mutants are listed in Table I, and the locations of the introduced mutations and relevant restriction sites are shown in Fig. 2. To facilitate construction of the K330A mutant, PCR mutagenesis was used to introduce a Smal site into PHR1 at nucleotides 684–869 (relative to the first PHR1 ATG; Fig. 2), yielding plasmid pGBS425. The amino acid sequence of photolyase was not changed in this construction. Alanine substitution mutations were introduced at Lys$^{330}$, Glu$^{384}$, Arg$^{452}$, Phe$^{494}$, and Gln$^{514}$ using two-primer (Lys$^{330}$, Arg$^{452}$) or four-primer PCR mutagenesis (Lys$^{384}$, Glu$^{384}$, Phe$^{494}$, and Gln$^{514}$) and pGBS424 as template. DNA polymerase (New England Biolabs) was used in the first PCR stage (25 cycles). For 4-primer PCR, the first stage PCR products were purified on 3% GTG agarose gels (FMC Bioproducts), recovered in sterile water, and used for the second PCR reaction in conjunction with appropriate outside primer pairs (Table I). Amplification was carried out for 15 cycles using Taq polymerase (Life Technologies, Inc.). Purified PCR products were digested with appropriate restriction endonucleases and cloned into pGBS425 which had been digested with the same set of enzymes as follows: Kpn1 and Xba1 for K330A and E384A, Xba1 and Pst1 for F494A and Q514A, Kpn1 and Xba1 for R452A, and Smal and Kpn1 for K330A (Fig. 2). In each case the nucleotide sequence of the amplified region was verified to ensure that no additional mutations were introduced by the PCR. The photolyase bearing three substitutions (K330A/E384A/F494A) was constructed by subcloning restriction fragments containing the E384A and the F494A mutations into the K330A plasmid using unique restriction sites surrounding each mutation. To express PHR1 without any attached fusion protein, the ~2.4-kilobase pair BglII-Pst1 fragment from each mutant was subcloned into similarly digested pCB12144/recon. Construction of mutant W387A has been described previously (9).

Photolyase Purification and Spectral Characterization—All photolyses were overexpressed from pCB12144/recon derivatives (no maltose binding protein fusion) in E. coli strain CR6036P'f'ac' (Phr') and purified as described previously to greater than 95% purity as judged by Coomassie Blue staining (9). The protein concentration and chromophore content for each protein preparation were determined by spectrophotometry from 220 to 700 nm (9). The spectra were closely examined for the presence of peaks near 450, 490, 580, and 625 nm, which are indicative of oxidation of the reduced flavin chromophore to the blue neutral radical or FAD$_{\alpha}$ (16).

Quantitation of Specific and Nonspecific Equilibrium Association Constants ($K_a$, $K_s$, $K_{NS}$)—A 43-bp DNA substrate containing a centrally located thymine dimer was prepared for these studies as shown in Oligonucleotide 1.

$$5'-\text{GCGTATCGATGGCCTGCAGGCAAGT} \rightarrow \text{TGGAGGATTTCGCTAGTG} \rightarrow \text{GCGTATCGATGGCCTGCAGGCAAGT} \rightarrow \text{TGGAGGATTTCGCTAGTG}$$

Oligonucleotide 1

A (cis,syn)-thymine dimer (a gift from Xiaodong Zhao and Aziz San-
car) was synthesized and incorporated into the top strand oligonucleotide shown using standard oligonucleotide coupling protocols (17). Full-length oligonucleotide was purified from a denaturing polyacrylamide gel, quantitated by absorbance at 254 nm, and the two oligonucleotides were mixed at a concentration of 500 nM dimer oligonucleotide and 1500 nM bottom strand oligonucleotide in 100 mM Tris (pH 8.0). A sample of 10 mM Tris (pH 8.0) with 1 mM EDTA was used as competitor for nonspecific binding studies.

For each polynucleotide preparation the fraction of polynucleotide molecules in DNA binding was determined by titrating a fixed concentration of enzyme (2–3 $K_d$) with increasing amounts of substrate, as described previously (9). Electrophoretic mobility shift assays (EMSA) were used to separate bound and free substrate. Radioactivity present in the bound and free DNA bands was quantitated using an Ambis Radioanalytic Imager (Ambis Systems). The equilibrium association constants ($K_a$) of the various photolyases for the dimer-containing substrate were determined by titrating substrate (5 $\times$ 10$^{-3}$ M) with enzyme (9). Control reactions without enzyme were used to calculate the amount of background smearing of the DNA in each gel. The slope of a Eadie-Scatchard plot ($S/K_a$ versus $E/K_a$) of these data yielded $K_a$ (where ES = bound DNA concentration; S = total substrate concentration; and $E/K_a$ = free enzyme concentration). The binding affinity for nonspecific DNA ($K_{NS}$) was measured by titrating the enzyme-dimer oligonucleotide complex (70% of substrate bound in the absence of competitor) with cold oligo(dT)$_{18}$ over the nucleotide concentration range of 1 $\times$ 10$^{-3}$ to 5 $\times$ 10$^{-3}$ M. The x-intercept of a plot of 1/ES versus $K_a$ yielded $-1/K_a$ (18). $K_a$, the intrinsic specific association constant of photolyase for the dimer, was determined from the relationship $K_{obs} = K_a(1 + [D]/K_{NS})$ where $D = $ molar concentration of nondimer nucleotides in the 43-bp substrate (9). Each $K_a$ was measured and $K_{NS}$ and active molecule determination was performed at least 3 times.

**Ethylnitrosourea Interference**—End-labeled double-stranded dimer-containing substrate was ethylated at phosphate groups (19) as follows: 100 $\mu$L of ethanol saturated with ethylnitrosourea (Sigma) were added to 4 pmol of labeled substrate dissolved in 100 $\mu$L of sodium cacodylate (pH 8.0). Following incubation at 50 °C for 1 h, ethylnitrosourea was removed by seven ethanol precipitations in 0.5 M CH$_3$CO$_2$NH$_4$ with 25 $\mu$g of carrier RNA. The DNA was washed with 95% ethanol after the final precipitation, dissolved in 0.1 mM EDTA, and recovered substrate was quantitated by scintillation counting.

Electrophoretic mobility shift assays, in a volume of 100 $\mu$L, were performed using the conditions described above. 5 $\times$ 10$^{-3}$ M substrate was incubated with sufficient photolyase to produce 60% binding as judged by EMSA. Following EMSA, the bound and free DNAs were cut out of the gel, and the DNA was eluted by overnight agitation in 0.5 M CH$_3$CO$_2$NH$_4$, 1.0% SDS, filtered through glass wool, extracted with phenol and ether, and precipitated in ethanol. Following a second
RESULTS

Protein Modeling—By using the crystal structure of E. coli photolyase (14) as a starting point, we modeled the structure of yeast Phr1 photolyase over the region encompassing residues 383 through 532 (E. coli photolyase residues 273–422; Figs. 1 and 3). The validity of this approach is supported both by the high degree of sequence conservation in these regions of the enzymes (50% identity with no gaps; Fig. 1 and see Ref. 15), which suggests a similar fold, and by the results obtained during the modeling. Of the 73 nonidentical residues replaced during modeling, only 6 residues (Phr1 residues 395, 401, 417, 418, 478, and 521) produced steric clashes when modeled using the torsion angles of the equivalent E. coli residue. Each of these residues could, however, be modeled using alternative standard torsion angles. Most of these residues are solvent-exposed and none lie near the proposed DNA binding surface or the flavin-binding site; therefore it is unlikely that an incorrect choice of torsion angle for these residues would affect either the overall accuracy of the model or the structure of either the DNA-binding site or the flavin-binding site. Overall the structures of the enzymes over the modeled region are highly similar to one another (Fig. 3), as well as to the recently solved Aspergillus nidulans photolyase structure (21). The latter enzyme exhibits 70% sequence identity to the E. coli photolyase within the modeled region (21).

Two important results relevant to this study emerged from the modeling of the yeast enzyme. (i) The structure of the flavin-binding site is conserved. Within the modeled region there are 6 amino acids that interact with FAD (14), 5 of which are identical to those found in the E. coli enzyme (Fig. 1). The single amino acid substitution (Trp338(E. coli) → Tyr448(S. cerevisiae)) places the Tyr448 side chain OH within H-bonding distance of the same FAD phosphate contacted by the ring nitrogen of Trp338 (data not shown). Of the seven FAD-contacting residues that lie outside of the modeled region (14), 5 are identical in both enzymes and one of the nonidentical amino acids (Arg239(E. coli) → Gly340(S. cerevisiae)) is predicted to make contact only through backbone substituents. The same substitution occurs in the A. nidulans photolyase structure where the FAD contact is conserved (21). Based upon the conservation of the cofactor-binding site, we conclude that the orientation of FADH− is likewise retained. (ii) The crucial features of the proposed DNA binding surface of the E. coli enzyme (14) are conserved in yeast photolyase (Fig. 3). This conservation is seen most strongly for residues lining the cavity leading from the surface to the flavin chromophore. The three substitutions, Lys383(S. cerevisiae) → Asn373(E. coli), Phe494(S. cerevisiae) → Trp384(E. coli), and Lys516(S. cerevisiae) → Arg226(E. coli), conserve the asymmetric distribution of polar and hydrophobic residues in the cavity. The band of positive electrostatic potential extending out from the cavity is also conserved and is augmented by several substitutions: Lys383(S. cerevisiae) → Asn273(E. coli), Lys359(S. cerevisiae) → Gln406(E. coli), and Gln505(S. cerevisiae) → Ala393(E. coli). The structural similarity of the proposed DNA binding surfaces of the yeast and E. coli enzymes is consistent with the fact that the footprints made by the enzymes on dimer-containing DNA are essentially identical (12).

Based upon the results of the modeling study, we selected Phe494, Gln406, Lys359, Lys383, Arg452, and Gln514 of PHR1 as targets for alanine substitution mutagenesis. The first three residues lie within the active site cavity and are predicted to interact with a pyrimidine dimer flipped into the cavity, but they do not appear to interact directly with FAD. This latter observation is crucial to the interpretation of DNA binding data because mutations that destabilize flavin binding usually lead to unfolding of the enzyme. 2 Lys383, Arg452, and Gln514 lie outside of the cavity and along the region of positive electrostatic potential on the proposed DNA binding surface (14). Based upon preliminary docking experiments (data not shown), these residues are likely to interact with the DNA flanking the dimer.

Effects of Ala Substitutions on Substrate Binding and Discrimination—Wild-type and alanine-substituted photolyses described above were purified and characterized by UV spectroscopy. Photolyase from the previously reported mutant...
Table II

| Photolyase       | Location of altered amino acid(s) | $K_A^{\times 10^6}$ M$^{-1}$ | $K_{NS}^{\times 10^5}$ | Discrimination ratio | Relative quantum yield at 365 nm$^d$ |
|------------------|-----------------------------------|-------------------------------|------------------------|----------------------|-------------------------------------|
| Wild-type        |                                   | 11.4 (±0.4)                   | 1.8                    | 6.3                  | 1.0                                 |
| K330A            | Inside cavity                     | 6.3 (±1.0)                    | 0.9                    | 7.2                  | 1.1                                 |
| E384A            | Inside cavity                     | 6.7 (±1.3)                    | 1.6                    | 4.1                  | 0.4                                 |
| F494A            | Inside cavity                     | 4.5 (±1.4)                    | 1.3                    | 3.5                  | 0.7                                 |
| K330A/E384A/F494A| Inside cavity                     | 0.9 (±0.1)                    | 0.8                    | 1.0                  | 0.2                                 |
| W387A            | Inside cavity                     | 0.7 (±0.1)                    | 0.6                    | 1.1                  | 0.2                                 |
| K383A            | Outside cavity                    | 1.7 (±0.2)                    | 1.0                    | 1.8                  | 1.0                                 |
| R452A            | Outside cavity                    | 0.6 (±0.1)                    | 3.1                    | 0.2                  | 0.4                                 |
| Q514A            | Outside cavity                    | 0.4 (±0.1)                    | 0.7                    | 0.6                  | 0.9                                 |

$^a$ Affinity for a 43-bp substrate containing a single pyrimidine dimer. Standard errors are indicated in parentheses.

$^b$ Affinity per mol of nucleotide in oligo(dT)$_{18}$.

$^c$ Discrimination ratio = $K_A/K_{NS}$. Under our experimental conditions, $K_A = K_{NS}(1 + -0.005)$, therefore $K_A$ was equated with $K_{NS}$ for this calculation.

$^d$ Relative to the quantum yield of the wild-type enzyme.

W387A (9) was also purified and used for comparative purposes in all of the studies described below. For all enzyme preparations, the folate and flavin chromophores were present in approximately equimolar (0.9–1.0) stoichiometry with the apoenzyme, and neither the flavin blue neutral radical nor oxidized flavin were detected (data not shown). Thus the overall structure of the enzyme was not perturbed in the mutants; the flavin-binding site was intact, and the normal oxidation state of the flavin chromophore was retained. The integrity of the flavin-binding site and retention of the normal oxidation state of the flavin chromophore are particularly noteworthy. The dinucleotide flip model predicts that the dimer interacts with both the adenine and isoalloxazine rings of flavin (14), and photolyase lacking flavin does not bind pyrimidine dimers with measurable affinity (22). Furthermore, although the redox state of flavin does not alter DNA binding, oxidation of the flavin chromophore to the blue neutral radical reduces the quantum yield of photolysis by an order of magnitude and enzyme containing oxidized FAD is inactive in photolysis (5, 22).

The equilibrium binding affinities of the photolyases for the 43-base pair substrate containing a single pyrimidine dimer were determined by EMSA. Each of the substituted photolyases displayed reduced affinity for the dimer (Table II). The $K_A$ values for enzymes bearing single substitutions in the active site cavity varied widely. Substitution at Lys330, Phe494, or Glu384 produced a 40–60% decrease in affinity, whereas substitution of Trp387 decreased binding 16-fold. These results suggest that individually Lys330, Phe494, and Glu384 contribute little to the overall binding energy, whereas Trp387 makes a major contribution. That these residues do indeed participate in binding was evident from the $K_A$ of the photolyase in which all three amino acids were substituted simultaneously with alanine. The reduction in binding affinity exhibited by the triple mutant (K330A/F494A/E384A) was comparable to that seen with the W387A mutant. Phe494 and Glu384 in particular are deeply recessed into the cavity and cannot contact any residue in normal B-DNA. In addition, both E. coli and yeast photolyase bind pyrimidine dimers in single-stranded DNA with an affinity similar to that seen with double-stranded DNA (6), which rules out binding to an extrahelical base in the complementary strand as a binding determinant. Therefore the reduced binding affinity exhibited by these mutants strongly supports the dinucleotide flip model for photolyase binding.

Substitutions outside of the active site cavity produced larger decreases in affinity (Table II). The $K_A$ for interaction between dimer-containing DNA and the K383A mutant was decreased approximately 7-fold, whereas $K_A$ values for R452A and Q514A decreased 21- and 29-fold, respectively. The large decrease in binding affinity seen with these mutants, compared with the smaller decreases seen with most of the cavity mutants, argues that much of the free energy of binding comes from interactions between photolyase and DNA flanking the dimer rather than from direct interaction with the dimer. This is consistent with previous observations suggesting that photolyase recognizes not only the dimer but also DNA structural components flanking the dimer (9, 12, 13).

The ability of a DNA-binding protein to recognize its specific target among an excess of nonspecific binding sites is determined by the ratio of the binding constants $K_A/K_{NS}$ (specific binding/nonspecific binding) known as the discrimination ratio. Among the photolyases bearing single substitutions in the cavity, only the W387A mutant displayed a large reduction in discrimination ratio (Table II). Once again synergy was observed with the triple mutant more severely compromised in substrate discrimination than any of the single cavity mutants including the W387A mutant. Two of the single substitutions outside of the cavity produced larger decreases in discrimination ratio than did any of the substitutions (single or multiple) inside the cavity (Table II). Particularly noteworthy is the R452A substitution which produced a 3.7-fold increase in $K_{NS}$ and 20-fold reduction in $K_A$. Among the 10 Phr1 substitution mutants now characterized (Ref. 9 and this work), this is the only mutant that exhibits an increase in $K_{NS}$. Clearly Arg452 is a major determinant of binding specificity.

Ethylation Interference Studies—Interactions between photolyase and DNA phosphates surrounding the dimer contribute to both specific and nonspecific substrate binding (9). Ethylation DNA phosphates interferes with these interactions either by eliminating a negative charge or by steric interference and is therefore a useful method for probing interactions at the binding interface. The results of ethylation interference studies on the mutant photolyases and wild-type enzyme are shown in Figs. 4 and 5. In previous studies we demonstrated that ethylation of the first phosphate 5'- to the dimer and the first through the fourth phosphates 3'- to the dimer interferes with binding by wild-type photolyase, whereas ethylation of the intradimer phosphate has no effect (9, 12). Interference at the fourth phosphate 3'- to the dimer is generally weak, and in the experiments reported here was not clearly discernible. In all other respects the ethylation interference pattern shown for wild-type photolyase in Fig. 4 is identical to those reported previously. In contrast, each of the mutants displayed changes in the ethylation interference pattern consistent with alterations in the binding interface.

Among the active site cavity mutants, two distinct patterns of altered ethylation interference were apparent (Fig. 5). Photolyases W387A and K330A exhibited decreased interference at
all sites, suggesting that interactions between the enzymes and phosphates in the dimer-containing strand have changed along the entire binding interface. Despite this general similarity, the two photolyases displayed distinctive differences in the interference pattern at specific phosphates indicating that the DNA repaired by photolyase was quantified by electrophoretic mobility shift assay. Under the experimental conditions employed, ≥70% of the substrate was bound prior to exposure to photoreactivating light. Thus multiple cycles of binding and photoreactivation by a single enzyme molecule do not contribute significantly to the results obtained.

The quantum yield value obtained for each photolyase relative to wild-type is shown in Table II. Among the substitutions within the cavity, K330A and F494A had little or no effect on the quantum yield. In contrast, E384A, the triple substitution mutant, and W387A displayed 60–80% reductions in quantum yield relative to the wild-type enzyme. The Glu384 side chain lies in the floor of the active site cavity (Fig. 6), and its negative charge may assist in directing electron transfer from FADH$^+$ to the dimer and away from solvent. The further reduction in the quantum yield seen with the R452A mutant resembled that of the triple mutant (K330A/E384A/F494A) suggesting that, despite its surface location, Arg452 plays a key role in orienting the dimer in the active site cavity. This interpretation is supported by the results of quantum yield experiments discussed below. The Q514A mutant displayed an unusual pattern in that interference increased at all sites. The simplest explanation is that one or more strong nonphosphate contacts have been lost in the mutant and that, as a result, interactions with phosphates contribute relatively more to the overall binding energy.

The Effects of Substitution Mutations on Quantum Yield—Substitutions that affect the electronic environment within the active site, the positioning of the dimer within the active site, or the equilibrium between the flipped and nonflipped state potentially alter the efficiency of photolysis. Therefore we determined the quantum yield for photolysis by each substituted enzyme and compared it to the quantum yield of the wild-type enzyme. Each of the purified proteins was bound to dimer-containing substrate, and the complex was photoreactivated with 365 nm light in increments of 25 J/m$^2$. The amount of DNA repaired by photolyase was quantified by electrophoretic mobility shift assay. Under the experimental conditions employed, ≥70% of the substrate was bound prior to exposure to limiting photoreactivating light. Thus multiple cycles of binding and photoreactivation by a single enzyme molecule do not contribute significantly to the results obtained.
A model of a pyrimidine dimer docked in the active site cavity of the yeast Phr1 photolyase is shown in Fig. 6. This model is based upon the crystal structure of a thymine dimer (23) and the modeled yeast photolyase structure; until the structure of a pyrimidine dimer in double-stranded DNA is solved at atomic resolution, this model provides a useful context for interpreting the results of this and previous structure-function studies (9, 24). The 5' → 3' placement of the dimer is predicated upon our observation that the enzyme interacts more extensively with DNA residues 3' to the dimer and upon the locations of mutations that reduce DNA binding (this work and Refs. 9, 12, and 13). In this model the 5' base in the dimer is involved in π-π stacking interactions with Trp387; the methyl group of the 5' base is sandwiched into a hydrophobic pocket between Trp387 and Phe494; the methyl group of the 3' base is involved in hydrophobic interactions with Phe494; N-3 of both bases are within hydrogen bonding distance of O-ε of Glu384, and the phosphate 5' to the dimer is within hydrogen bonding distance of the side chains of Lys330 and Lys383. Other potential interactions involving the dimer include a hydrogen bond between the exocyclic amine of the adenine base in FAD and O-4' of the 5' dimer, and a "hydrogen bond" (25) between the sulfur atom of the 5' base and Trp387 side chain simultaneously lower the energetic cost of placing of the dimer is shown in red. To show more clearly the relationship of the dimer bases to residues in the cavity, the structure has been rotated approximately −25° around the y axis relative to the view shown in Fig. 3. The entire modeled region is shown.

of the active site cavity, and thus alanine substitution was not expected to affect the electronic environment within the active site. Indeed, none of the other substitutions outside of the cavity significantly changed the quantum yield (Table II). An attractive explanation is that interaction between Arg452 and DNA residues flanking the dimer plays a crucial role in stabilizing the dimer in the flipped state.

**DISCUSSION**

**A Model for Dimer Binding by Photolyase**—Approximately 50% of the total energy of enzyme-substrate complex formation between photolyase and dimer-containing DNA comes from interaction between the enzyme and elements of the dimer (7, 12, 13). If, as proposed by Park et al. (14), the dimer resides within the cavity leading from the flavin chromophore to the surface of the enzyme, interactions with residues lining the cavity should contribute significantly to the binding energy. The results reported here provide direct experimental confirmation of this prediction. Alanine substitution at each of four sites (Lys330, Glu384, Phe494, and Trp387) within the active site cavity reduces substrate binding. Because these residues lie too deeply buried in the pocket to interact with normal B-DNA, the only explanation for these results is that either the enzyme or the substrate undergoes a dramatic structural alteration that places these amino acids within interacting distance with DNA. A large scale conformational change in the enzyme is highly unlikely. With the exception of Lys330, all of the cavity residues probed here lie within helices that are held firmly in place by multiple interactions with neighboring structural elements, and movement of these helices toward the surface would disrupt multiple interactions with the flavin chromophore. These considerations, combined with the structural and chemical complementarity of the active site to the dimer, provide strong evidence that the cavity in photolyase identified by Park et al. (14) is indeed the binding site for an extrahelical pyrimidine dimer.
the helix. It should be noted that most of the interactions in the active site cavity can also occur when two nondimerized pyrimidines are flipped into the helix. This suggests that much of the binding specificity conferred by interactions within the active site cavity arises from differences in the energy cost of flipping two unlinked bases versus a dimer. Several studies suggest that the presence of a dimer in a DNA helix weakens stacking interactions with adjacent bases and hydrogen bonding with the partners of the dimerized bases (27–29), and thus the energetic cost of flipping a dimer out of the helix and into the active site cavity should be lower than the cost of flipping two noncovalently linked pyrimidines. The free energy contributed by interactions between cavity residues and the dimer is sufficient to stabilize the dimer in the flipped state but is not sufficient to stabilize two noncovalently linked pyrimidine bases flipped out of the helix. This line of reasoning suggests that the W387A mutant and the triple mutant (K330A/E384A/F494A) the equilibrium between the flipped and nonflipped dimer is shifted. Presumably flipping of the dimer is accompanied by repositioning of phosphates immediately surrounding the dimer, as has been seen for other flipped nucleotides (30, 31). Both the triple mutant and the W387A mutant display large decreases in ethylation interference at phosphates immediately surrounding the dimer. These are characteristics shared with the Trp455Trp mutant and the triple mutant (K330A/E384A/F494A) and suggest that, like these mutants, Arg452 may play an important role in maintaining the dimer in the flipped orientation. We note that several DNA-binding proteins that “flip” nucleotides out of the helix do so by inserting one or more side chains into the helix, thereby displacing the base (31, 33). Whether this is the role of Arg452 must await the co-crystal structure.

Evolutionary Implications—Two classes of CPD photolyases have been identified based upon sequence homology (34). Most microbial photolyases, including the E. coli, A. nidulans, and S. cerevisiae enzymes, are members of class I and share 25–43% sequence homology, whereas most photolyases from higher eukaryotes are members of class II and share 38–72% homology. Because the homology between enzymes in different classes is only 10–17%, it is pertinent to ask whether enzymes in these two classes employ common mechanisms to recognize and repair pyrimidine dimers. While direct information on the three-dimensional structure of class II enzymes is lacking, several lines of evidence support a common mechanism. All of the enzymes require reduced FAD to carry out photolysis (4, 34). The M. thermoautotrophicum photolyase, the only representative of class II enzymes that has been characterized extensively with respect to DNA binding, contacts the same phosphates surrounding the dimer as do the yeast and E. coli enzymes (35).

Finally, despite the low level of overall homology between the two classes of enzymes, there is striking conservation of amino acids lining the active site cavity, and to a lesser extent of amino acids that interact with flanking DNA residues (Table III). Even more surprising is the sequence conservation between residues in the active site cavity of the class I photolyases and the (6-4) photolyases (Table III). Evidence has been presented suggesting that the (6-4) photolyases also flip their substrates out of the DNA helix (36). However, the structure of the binding site must be sufficiently different to exclude pyrimidine dimers that are not efficiently bound by (6-4) photolyases (2, 37). Nevertheless, of the five residues within the cavity that are predicted to contact the dimer in CPD photolyases, four are conserved in the (6-4) photolyases. This is consistent with the proposed evolution of the class I CPD and (6-4) photolyases from a common ancestral gene (38, 39) and further suggests that the different binding specificities of the enzymes entail surprisingly few changes in the active site cavity. The co-crystal structures of these two types of photolyases should

### Table III

| Location and residue number in S. cerevisiae Phr1 photolyase | Residue found at equivalent position in class of photolyase |
|-------------------------------------------------------------|-------------------------------------------------------------|
| Inside active site cavity                                   |                                             |
| 330 Arg, Lys, His                                           | Arg, Lys, Asn                                             |
| 384 Glu, Gln                                                | Glu, Gln                                                   |
| 387 Trp                                                     | Arg, Lys, Asn                                             |
| 455 Met                                                     | Met, Lys, His                                             |
| 494 Trp, Phe                                                | Trp, Lys, His                                             |
| Outside active site cavity                                  |                                             |
| 383 Gly, Gln, Arg, Asn, Lys, Ser                            | Arg, Glu, Lys, Asn                                        |
| 452 Arg                                                     | Arg, Lys                                                   |
| 507 Arg                                                     | Arg, Lys                                                   |
| 514 Gin                                                     | Gin, Thr                                                   |
| 517 Arg, Lys                                                 | Arg, Lys                                                   |

* Based upon the alignment of the class I and class II photolyases in Ref. 34.

Based upon alignment of the predicted amino acid sequences of the 6–4 photolyases from A. thaliana (38) and D. melanogaster (39) using the alignment program CLUSTAL W.
provide insight into the crucial interactions required for such discrimination.

Acknowledgments—We thank Dr. Hengming Ke for assistance with the protein modeling; Dr. Hee-Won Park for providing the structure of the pyrimidine dimer; Drs. Aziz Sancar and Xiaodong Zhao for dimer substrate and useful discussions; and Kalpana Kasala for technical assistance.

REFERENCES
1. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
2. Todo, T., Takemori, H., Ryo, H., Ihara, M., Matsunaga, T., Nikaido, O., Sato, K., and Nomura, T. (1993) Nature 361, 371–374
3. Brash, D. E., Franklin, W. A., Sancar, G. B., Sancar, A., and Haseltine, W. A. (1985) J. Biol. Chem. 260, 11438–11441
4. Sancar, A. (1994) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
5. Todo, T., Takemori, H., Ryo, H., Ihara, M., Matsunaga, T., Nikaido, O., Sato, K., and Nomura, T. (1993) Nature 361, 371–374
6. Brash, D. E., Franklin, W. A., Sancar, G. B., Sancar, A., and Haseltine, W. A. (1985) J. Biol. Chem. 260, 11438–11441
7. Sancar, A. (1994) Biochemistry 33, 2–9
8. Todo, T., Ryo, H., Yamamoto, K., Toh, H., Inui, T., Ayaki, H., Nomura, T., and Ikenaga, M. (1996) Science 272, 109–112
9. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
10. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
11. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
12. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
13. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
14. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
15. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
16. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
17. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
18. Sancar, A. (1996) in Annual Review of Biochemistry (Richardson, C. C., Abelson, J. N., and Raetz, C. R. H., eds) pp. 43–81, Annual Reviews, Inc., Palo Alto, CA
19. Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980) Cell 20, 269–281
20. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499–560
21. Tamada, T., Hitokoto, K., Higuchi, Y., Inaka, K., Yasui, A., de Ruiter, P. E., Eker, A. P. M., and Miki, K. (1997) Nat. Struct. Biol. 4, 887–891
22. Payne, G., Wills, M., Walsh, C. T., and Sancar, A. (1990) Biochemistry 29, 5706–5711
23. Cadet, J., Vuituriez, L., Hruška, F. E., and Grand, A. (1985) Biopolymers 24, 897–903
24. Kim, S.-T., Li, Y. F., and Sancar, A. (1992)Proc. Natl. Acad. Sci. U. S. A. 89, 900–904
25. O’Gara, M., Klimasauskas, S., Roberts, R. J., and Cheng, X. (1996) J. Mol. Biol. 261, 634–645
26. Patrick, M. H., and Rahn, R. O. (1976) in Photochemistry and Photobiology of Nucleic Acids (Wang, S. Y., ed) pp. 35–95, Academic Press, New York
27. Pearlman, D. A., Holbrook, S. R., Pirkle, D. H., and Kim, S. H. (1985) Science 227, 1304–1308
28. Boyde, S., Stellman, S., and Hingerty, B. (1980) Biopolymers 19, 1695–1701
29. Haynes, P. N., Williams, D. L., Ratiff, R. L., Yarghese, A. J., and Rupert, C. S. (1971) J. Am. Chem. Soc. 93, 4940–4943
30. Vassylyev, D. G., Kashiwagi, T., Mikami, Y., Ariyoshi, M., Iwai, S., Ohtsuka, E., and Morikawa, K. (1995) Cell 83, 773–782
31. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokkan, H. E., and Trainer, J. A. (1996) Nature 384, 87–92
32. Taylor, J. S., Garrett, D. S., Brockie, I. R., Svoboda, D. L., and Telser, J. (1990) Biochemistry 29, 8858–8866
33. Reinsch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. (1995) Cell 82, 143–153
34. Yasui, A., Eker, A. P. M., Yasuhira, S., Yajima, H., Koyayashi, T., Takao, M., and Ohkawa, A. (1994) EMBO J. 13, 6141–6151
35. Kiener, A., Husain, I., Sancar, A., and Walsh, C. T. (1989) J. Biol. Chem. 264, 13880–13887
36. Zhao, X., Hsu, D. S., Zhao, S., Taylor, J.-S., and Sancar, A. (1997) J. Biol. Chem. 272, 32580–32590
37. Uchida, N., Mitani, H., Todo, T., Ikenaga, M., and Shimah, A. (1997) Photochem. Photobiol. 65, 964–968
38. Nakajima, S., Sugiyama, M., Iwai, S., Hitomi, K., Ootshi, E., Kim, S.-T., Jiang, C.-Z., Todo, T., Britt, A. B., and Yamamoto, K. (1998) Nucleic Acids Res. 26, 638–644
39. Todo, T., Ryo, H., Yamamoto, K., Toh, H., Inui, T., Ayaki, H., Nomura, T., and Ikenaga, M. (1996) Science 272, 159–112