Published online 20 May 2019

Nucleic Acids Research, 2019, Vol. 47, No. 12 ***6015–6028***
doi: 10.1093/nar/gkz359

NAR Breakthrough Article

Structure and mechanism of pyrimidine–pyrimidone (6-4) photoproduct recognition by the Rad4/XPC nucleotide excision repair complex

Debamita Paul[1,†], Hong Mu[2,†], Hong Zhao[3], Ouathek Ouerfelli[3], Philip D. Jeffrey[4], Suse Broyde[2,*] and Jung-Hyun Min[1,*]

[1]Department of Chemistry & Biochemistry, Baylor University, Waco, TX 76798, USA, [2]Department of Biology, New York University, New York, NY 10003, USA, [3]Organic Synthesis Core, Chemical Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA and [4]Department of Molecular Biology, Princeton University, NJ 08544, USA

Received March 11, 2019; Revised April 22, 2019; Editorial Decision April 25, 2019; Accepted April 26, 2019

ABSTRACT

Failure in repairing ultraviolet radiation-induced DNA damage can lead to mutations and cancer. Among UV-lesions, the pyrimidine–pyrimidone (6-4) photoproduct (6-4PP) is removed from the genome much faster than the cyclobutane pyrimidine dimer (CPD), owing to the more efficient recognition of 6-4PP by XPC-RAD23B, a key initiator of global-genome nucleotide excision repair (NER). Here, we report a crystal structure of a Rad4–Rad23 (yeast XPC-Rad23B ortholog) bound to 6-4PP-containing DNA and 4-μs molecular dynamics (MD) simulations examining the initial binding of Rad4 to 6-4PP or CPD. This first structure of Rad4/XPC bound to a physiological substrate with matched DNA sequence shows that Rad4 flips out both 6-4PP-containing nucleotide pairs, forming an 'open' conformation. The MD trajectories detail how Rad4/XPC initiates 'opening' 6-4PP: Rad4 initially engages BHD2 to bend/untwist DNA from the minor groove, leading to unstacking and extrusion of the 6-4PP:AA nucleotide pairs towards the major groove. The 5′ partner adenine first flips out and is captured by a BHD2/3 groove, while the 3′ adenine extrudes episodically, facilitating ensuing insertion of the BHD3 β-hairpin to open DNA as in the crystal structure. However, CPD resists such Rad4-induced structural distortions. Untwisting/bending from the minor groove may be a common way to interrogate DNA in NER.

INTRODUCTION

Ultraviolet (UV) radiation from sunlight is a ubiquitous and potent DNA-damaging mutagen (1–4). When DNA is exposed to UV, the most frequent lesions are covalent linkages between two adjacent pyrimidines, notably cyclobutane pyrimidine dimer (CPD) and 6-4 photoproduct (6-4PP). Failure in detecting and repairing such DNA damage is a major step leading to mutagenesis and cancer (5–9). The yields of CPD and 6-4PP upon UV irradiation vary depending on DNA sequence and also on UV wavelengths (1,10). In general, CPD is generated 3–4 times more than 6-4PP with UV-C and UV-B radiation at $\leq$296 nm and is a major source of mutations in mammalian cells (11). Nevertheless, 6-4PP is also cytotoxic and mutagenic although the high mutability of 6-4PP is likely suppressed by rapid repair of the lesion in cells (1,4,12,13). The biological impact of UV lesions, especially that of 6-4PP, may increase as stratospheric ozone that serves as a protective barrier against UV-C continues to decline (14).

In eukaryotes, the evolutionarily conserved, nucleotide excision repair (NER) pathway is key to removing the UV-lesions. Impaired NER genes decrease cell survival after UV in yeast and other cell lines and can cause the xeroderma pigmentosum (XP) cancer predisposition syndrome in humans, marked by extreme sun sensitivity and >1000-fold higher risk of sunlight-induced skin cancers (1,15,16). NER

*To whom correspondence should be addressed. Tel: +1 254 710 2095; Email: JungHyun_Min@baylor.edu
Correspondence may also be addressed to Suse Broyde. Tel: +1 212 998 8231; Email: broyde@nyu.edu
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Present address: Hong Zhao, Department of Chemistry, New York University, New York, NY 10003, USA.

can operate via two subpathways depending on how the lesions are initially recognized (6,17). For the lesions on the DNA strands that are being actively transcribed by RNA polymerases, an RNA polymerase stalled at a lesion signals the presence of the lesion and initiates the transcription-coupled NER (TCR) (18–21). In TCR, the repair rates of CPD and 6-4PP are similar to each other and are generally efficient (22,23). On the other hand, the global genome NER (GGR) can repair lesions on any location in the genome and relies on dedicated lesion recognition factors such as the UV-DDB complex (containing the DDB2/XPE protein) and the XPC-RAD23B-CETN2 complex (8,24–27). In this pathway, the recognition and repair rates of CPD and 6-4PP are significantly different. CPD is poorly recognized by XPC alone but requires UV-DDB. The UV-DDB ubiquitin ligase complex can bind to CPD within chromatin and help XPC to localize to the lesion (28–32). However, even in the presence of UV-DDB, the CPD repair is sluggish taking ∼24 hrs to remove 50% of CPD in cells (28,32–34). In contrast, 6-4PP is efficiently recognized by XPC and rapidly repaired within hours, even without UV-DDB (35–39). Once a lesion is recognized, both TC- and GG-NER proceed through a common pathway where the transcription factor IIH complex (TFIIH) is recruited to the lesion. TFIIH, containing XPD and XPB helicases, in turn verifies the presence of a bulky lesion and recruits other subsequent factors including XPA and RPA (26,40–43). Successful lesion recognition and proofreading lead to the excision of the lesion-containing single-stranded DNA (24–32 nucleotides) by XPG and XPF-ERCC1 endonucleases, followed by repair synthesis and nick sealing by DNA ligases (9,44).

The lesion recognition involving XPC is an indispensable, rate-limiting step in GG-NER and XPC is strictly required for the recruitment of TFIIH (8,24–26,37,44–49). The importance of XPC in human health and cancers has been well documented (50–55). *In vitro*, the XPC-RAD23B heterodimeric complex is necessary and sufficient to bind specifically to 6-4PP and other bulky lesions repaired by NER such as adducts derived from polycyclic aromatic hydrocarbons (e.g. benzo[*a*]pyrene), aromatic amines (e.g. acetylaminofluorene) and cisplatin (9,35,37,56–59). Notably, the binding specificities and NER excision efficiencies for lesions vary widely: various factors including conformation in DNA, lesion topology, stereochemistry, nature of the adducted base and sequence context play a role (35,37,56–65). For many lesions, the more destabilizing and distorting a lesion is, the better it is recognized and repaired. For instance, 6-4PP-harboring DNA exhibits greater distortions and is more dynamic than DNA containing CPD and it is a better substrate for XPC and NER (66–72). Although CPD is a poor substrate in its natural, matched sequence context, its recognition and repair efficiency *in vitro* dramatically improves if CPD is placed within a string of multiple mismatches such as TTT/TTT (36,38,39,45).

While a high resolution structure of a mammalian XPC-Rad23B is still lacking, crystal structures of its yeast ortholog, Rad4–Rad23 bound to DNA model lesions (a TTT/TTT mismatch bubble and a TTT/TTT enclosing a CPD lesion) have been solved (45). Being evolutionarily conserved from yeast to humans, the lesion recognition properties of Rad4 and human XPC are highly similar (73,74). Both XPC and Rad4 bind to mismatch bubble DNA with biochemical specificities comparable to those of *bona fide* NER lesions (Figure 1) (36,45). A recent cryo-EM architecture of the human XPC complex, albeit at low resolution of ∼25 Å, also indicates that the overall architecture of human XPC would be consistent with the crystal structure of yeast Rad4 (74). The crystal structures of Rad4 bound to these model lesions showed that the binding caused two nucleotide pairs harboring the mismatches (and CPD) to be flipped out of the DNA duplex and a β-hairpin from the BHD3 domain was inserted into the DNA duplex to fill the gap. Notably, in this 'open' structure, Rad4 did not directly contact the flipped-out CPD nucleotides but interacted exclusively with the nucleotides flipped out from the undamaged, complementary strand. The third mismatched base pair remained intrahelical. The structures therefore supported an indirect mode of recognition that helps explain the wide substrate specificity of the protein (75–78).

More recently, Min and Ansari have also observed that Rad4 can form the same 'open' structure when chemically crosslinked to normal DNA. Based on this and the kinetics of Rad4-induced DNA 'opening' as determined by temperature-jump perturbation spectroscopy (T-jump) (79), a 'kinetic gating' model has been proposed. In this 'kinetic gating' mechanism, the Rad4/XPC-binding specificity (recognition) is determined by the kinetic competition between the protein either opening the DNA site or diffusing away (79). At a lesion, the time required to form an open complex must be small enough to be accomplished while the protein resides on a given DNA register. It was further proposed that distorting and destabilizing lesions decrease the opening time by lowering the free energy barrier for DNA opening while it would also progressively increase the residence time of Rad4/XPC.

These studies further point to the importance of understanding the structural mechanism and the process by which the lesion search and 'opening' is carried out. Subsequent work using T-jump indicated that Rad4 can dynamically untwist DNA for nonspecific interrogation and that this untwisting does not require BHD3 (80). Studies by Van Houten and Min using atomic force microscopy and single-molecule fluorescence microscopy also showed that the mutant Rad4 lacking BHD3 can bend DNA on nonspecific as well as specific sites and that the mutant also localizes to the lesion sites such as fluorescein-dT, similarly as the wild type protein does (81). These results together suggested that Rad4 may use energetically coupled DNA untwisting and bending as it searches along and 'opens' a damaged site, for which BHD3 can be dispensable. This paper also showed that Rad4 shows anomalous diffusion around CPD, not allowing the protein to stay stationary enough to open the lesion site. Interestingly, recent computational MD studies by Mu *et al.* from the Broyde group showed that Rad4 engages its BHD2 domain early in the binding process for various high-specificity, polycyclic aromatic hydrocarbon adduct lesions indicating a role for BHD2 in facilitating BHD3 hairpin insertion into the DNA in a later step (82).

Despite these studies, a structure of Rad4/XPC bound to a *bona fide* substrate such as 6-4PP that is efficiently recognizable in its natural, matched DNA sequence context has been conspicuously missing. This has been a sore spot in

the field that limited the validation of the current mechanistic models. To fill the gap, we have solved the crystal structure of the Rad4–Rad23 complex bound to a 6-4PP lesion. The structure is congruent with other previously solved Rad4–DNA structures and additionally showed insight into how the protein may reach the 'open' conformation. Subsequently, we examined the process of initial Rad4–DNA interactions using molecular dynamics (MD) simulations. The results reveal a sequence of intricate conformational rearrangements that the Rad4–6-4PP DNA complex undergoes during the recognition. The process involved extensive engagement of BHD2 with the minor groove of the DNA, which facilitates DNA untwisting, bending, and nucleotide flipping towards the major groove, prior to the full 'opening'. We also show how these dynamical processes differ for the poorly recognized CPD. Altogether, our study presents a 3D structural trajectory for Rad4/XPC's photolesion binding that helps us understand the lesion recognition function of Rad4/XPC in an unprecedented level.

MATERIALS AND METHODS

Synthesis of 6-4 photoproduct-containing phosphoramidites

A batch of thymidine dimer methyl ester was irradiated with UV ($\lambda = 254$ nm) to generate UV-induced intra-strand crosslink adducts including 6-4PP. The resulting mixture of UV lesions and unmodified thymidine dimers was separated over a reverse-phase HPLC column (C-18 XBridge, 5 μm, 19 × 150 mm, Waters). The isolated 6-4PP diol was subsequently functionalized and purified over HPLC to generate 5′-dimethoxytrityl-6-4PP-3′-phosphoramidite suitable for automated oligonucleotide synthesis. NMR spectra were consistent with those reported previously (83). See SI Methods for more details.

Oligonucleotide synthesis

Oligonucleotides were synthesized by automated systems using phosphoramidite chemistry by IDT or MWG and were purified by HPLC. The sequences of the DNA are in Figures 1A and 2C. All oligonucleotides appeared as a single band on denaturing polyacrylamide gels as well as mass spectrometry. Annealing to make duplex DNA was done by slow cooling a 1:1 mixture of two complementary DNA strands from 95°C to room temperature over 5–8 h in 10 mM Tris–HCl, 1 mM EDTA, pH 8.0.

Preparation of Rad4–Rad23 complex

The Rad4–Rad23 and Rad4–Rad23–DNA complexes were prepared as previously described (45). Briefly, the Hi5 insect cells co-expressing the Rad4–Rad23 complex were harvested 2 days after infection. After lysis, the proteins were purified using His-Select Nickel agarose resin (Sigma) and anion exchange chromatography (Source Q, GE Healthcare), followed by thrombin digestion and cation exchange (Source S, GE Healthcare) and gel-filtration (Superdex200, GE Healthcare) chromatography. The final sample was concentrated by ultrafiltration to ∼13 mg ml^{-1} in 5 mM bis–tris propane–HCl (BTP-HCl), 800 mM NaCl, 5 mM dithiothreitol (DTT), pH 6.8.

Characterization of Rad4–DNA binding and DNA duplex thermal stabilities

The apparent binding affinities ($K_{d,app}$) were determined by competition electrophoretic mobility shift assays carried out in EMSA buffer (5 mM BTP-HCl, 75 mM NaCl, 5 mM DTT, 5% glycerol, 0.74 mM CHAPS, 500 μg ml^{-1} BSA, pH 6.8) as previously described (45,80). The thermal stabilities of the DNA duplexes were measured as previously described (80). Details are also included in SI Methods.

Crystallization, structure determination and refinement

All crystals were grown by the hanging-drop vapor diffusion method at 4°C, mixing 1 μl of protein solution and 1 μl of crystallization buffer. Crystals of the complex appeared after a few days at 4 °C in wells containing 50 mM BTP-HCl, 100 mM NaCl, 14% (v/v) 1-propanol, 5 mM spermidine-HCl and 5 mM dithiothreitol, pH 6.8. The crystals were then harvested with a harvest buffer using 20–30% (w/v) polyethylene glycol (PEG) 200 or PEG400 as cryoprotectants and were subsequently flash-frozen in liquid nitrogen. Diffraction data were collected at −170°C and were processed with the HKL2000 suite (84) and XDS (85). The structure of the Rad4–Rad23–DNA complex was determined by molecular replacement method using the previous structure (PDB code 2QSH, Chains A and X containing only Rad4 and Rad23) as the search model and refined through multiple rounds of refinement in Phenix (Supplementary Table S1) (45,86,87). The final model contains residues 126–514 and 525–632 of Rad4, and 256–308 of Rad23. The coordinates have been deposited with PDB code 6CFI. Figures were generated by PyMOL (88).

Molecular modeling and molecular dynamics (MD) simulations

We used the AMBER16 package (89) with ff14SB force field (90), explicit water and counterions for MD simulations, and Discovery Studio (*Dassault Systèmes* BIOVIA) for molecular modeling. The structures along each trajectory were clustered using the principal component analysis (PCA) method in the Bio3D package (91). The best representative structure for each cluster is defined as the one frame that has the shortest RMSD for the heavy atoms of the lesion-containing 6-mer and the protein backbone atoms of BHD2 to all other frames. Full details concerning the force field, molecular modeling, MD simulation protocols and analyses are given in SI Methods and Supplementary Table S2.

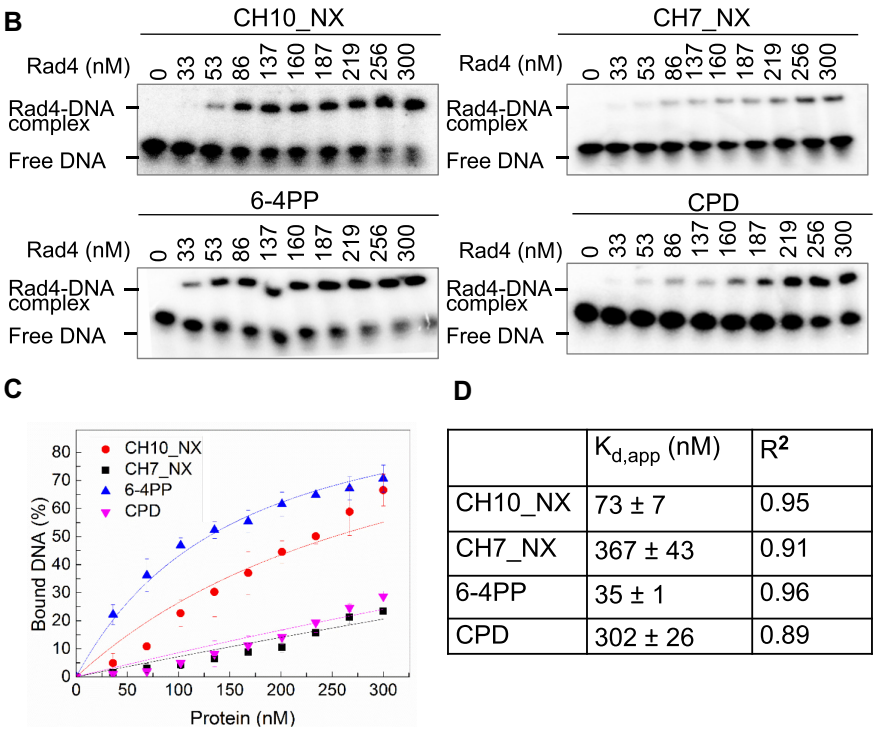
RESULTS

Differential binding of 6-4PP and CPD to Rad4 characterized by competitive gel-shift assays

Previous studies have established that human XPC binds to 6-4PP more specifically than to CPD *in vitro* (36,37) and that this correlates well with the greater relative repair efficiencies of 6-4PP compared with CPD in human cells and in CHO cells (92). The repair kinetics of CPD and 6-4PP in yeast cells show similar trends as with mammalian cells

A

DNA	Sequences	Tm (°C)
Mismatch (CH10_NX)	5′-TTGACTCGACATCC**CCC**GCTACAA -3′ 3′- ACTGAGCTGTAGG**CCC**CGATGTTA -5′	61.0
Normal, matched (CH7_NX)	5′-TTGACTCGACATCC**CCC**GCTACAA -3′ 3′- ACTGAGCTGTAGG**GGG**CGATGTTA -5′	76.5
6-4PP	5′-TTGACTCAACATCC**AAA**GCTACAA -3′ 3′- ACTGAGTTGTAGG**T**XXCGATGTTA -5′	61.3
CPD	5′-TTGACTCAACATCC**AAA**GCTACAA -3′ 3′- ACTGAGTTGTAGG**T**XXCGATGTTA -5′	69.2



D

	$K_{d,app}$ (nM)	R^2
CH10_NX	73 ± 7	0.95
CH7_NX	367 ± 43	0.91
6-4PP	35 ± 1	0.96
CPD	302 ± 26	0.89

Figure 1. Rad4–Rad23-binding affinities of 6-4PP- and CPD-containing DNA duplexes measured by competition gel-shift assays. (**A**) The construct names, DNA sequences and the melting temperatures (T_m) of the 24-bp DNA duplexes used as Rad4 substrates in the competition gel-shift experiments. The CCC/CCC mismatches used as a model lesion or the corresponding 3-bp positions in other DNA sequences are indicated in bold. XX denotes the position of 6-4PP or CPD thymidine-thymidine photodimers. (**B**) Typical gel images from the competition gel-shift assays for different DNA constructs. (**C**) Rad4-bound DNA fractions quantified from gels versus total protein concentrations. The symbols and error bars indicate the means and standard deviations, respectively, from triplicate EMSA experiments. Solid lines indicate the fit curves of the data points of the same color. (**D**) Apparent dissociation constant ($K_{d,app}$) and R^2 of the fits were calculated from the data points in (C). The errors in $K_{d,app}$ indicate the errors of the nonlinear regression fit.

(93,94), and it has been previously shown that the purified Rad4–Rad23 complex can bind to UV-irradiated DNA (95,96) and that enzymatic removal of CPD in the DNA did not diminish the binding of the UV-damaged DNA to Rad4, indicating a preference towards 6-4PP (95). However, the differences between 6-4PP and CPD for their binding to Rad4 have never been directly quantified with synthetic substrates. To fill this gap, we have synthetically prepared CPD- and 6-4PP-containing DNA duplexes. The melting temperature measurements of duplex DNA containing each lesion indicate that 6-4PP is more destabilizing than CPD (Figure 1A and Supplementary Figure S1). We then carried out competitive electrophoretic mobility shift assays (EMSA or gel-shift assays) using these DNA duplexes as ^{32}P-labeled substrates in the presence of a defined, matched DNA duplex (CH7_NX) as a nonspecific-binding competitor (79,80,97) (Figure 1). The apparent dissociation constant ($K_{d,app}$) for 6-4PP was ∼9-fold lower than that for CPD (34.5 ± 1.1 versus 302 ± 26) which binds with almost similar affinity to an undamaged control DNA (Figure 1). Intriguingly, 6-4PP bound ∼2-fold tighter than the 3-bp CCC/CCC mismatch sequence which has been thus far the best substrate that we have tested among different mismatched or CPD/mismatched DNA constructs (45,97). The results thus verify that the UV-lesion binding characteristics are indeed similar between human XPC and yeast Rad4, recapitulating the previous findings for their similarities in binding to other bulky DNA adduct substrates (73).

Rad4–Rad23 bound to 6-4 photoproduct forms an 'open' structure

Using the same 6-4PP DNA used in EMSA, we were also able to solve a 3.3 Å crystal structure of the Rad4 bound to the DNA (Supplementary Table S1). The overall structure of the Rad4–Rad23 complex bound to the 6-4PP-

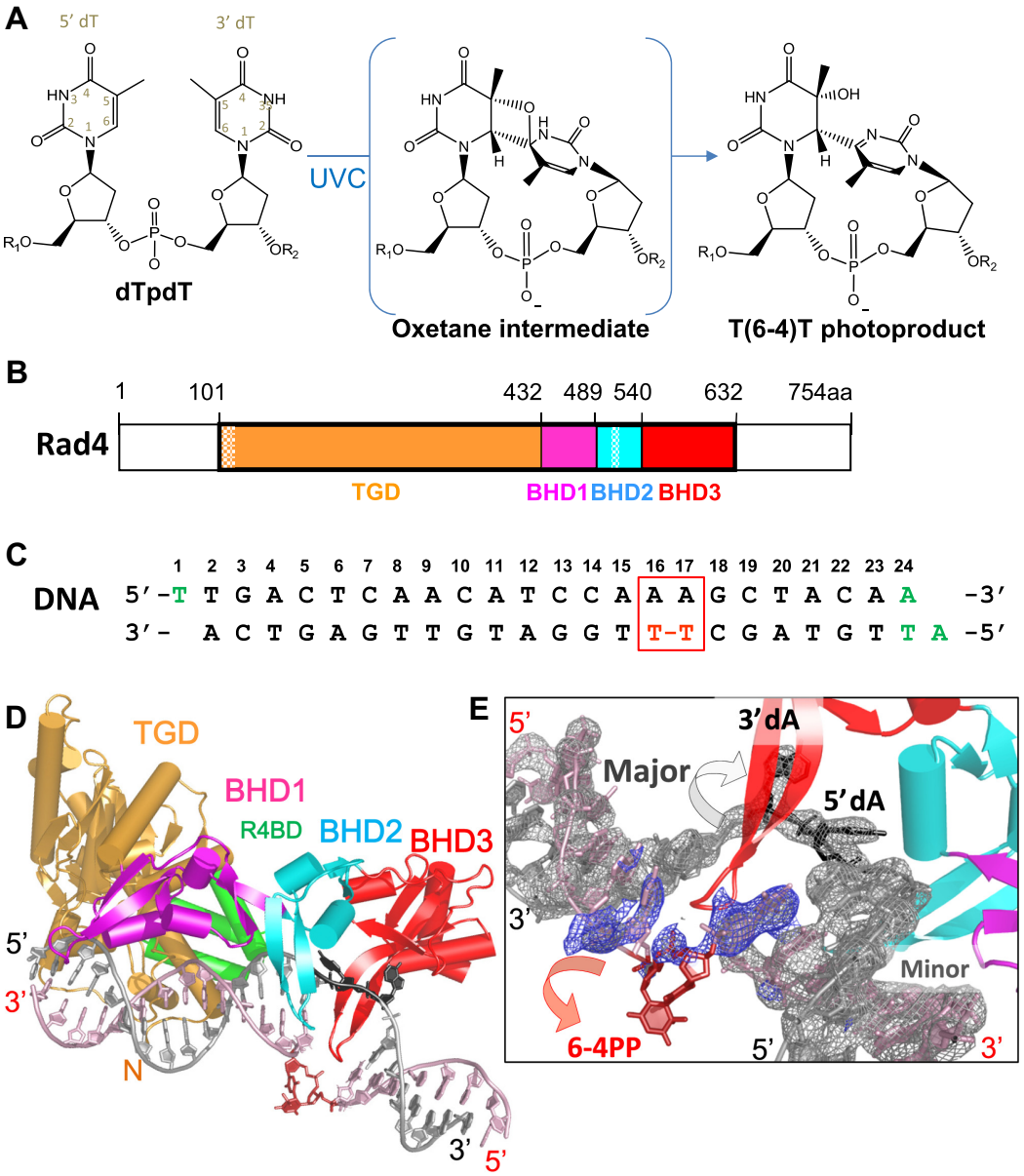


Figure 2. Structures of the Rad4–Rad23 complex bound to 6-4PP-containing DNA duplex. (**A**) Pyrimidine-pyrimidone (6-4) photoproduct (6-4PP) is induced by ultraviolet radiation of dipyrimidine nucleotide through an oxetane intermediate. Chemical structures are shown for the reaction between two thymines in a thymidine-thymidine dinucleotide (dTpT). (**B**) Domain arrangements and boundaries of Rad4 used in this study. The transglutaminase domain (TGD) is colored orange, β-hairpin domain 1 (BHD1) magenta, BHD2 cyan and BHD3 red. The crystallized Rad4 construct spans residues 101–632 as before (45). The disordered regions (residues 101–128, 518–525) in the crystal are checkered. Rad23 construct is the same as in (45). (**C**) Sequence of the 6-4PP DNA used in the co-crystallization. 6-4PP is indicated as T–T in red and the end sequences altered for MD simulation are in green (see Supplementary Figure S4). The red square indicates the four nucleotides that were flipped out by Rad4-binding in this structure. (**D**) Crystal structure of Rad4–Rad23 bound to 6-4PP-containing DNA duplex (PDB code 6CFI). The TGD (orange) and BHD1(magenta) of Rad4 bind to an 11-bp duplex segment of the DNA while BHD2 (cyan) and BHD3 (red) of Rad4 bind to a 4-bp segment in which two 6-4PP-containing nucleotide pairs are flipped out. The undamaged DNA strand is colored in silver while the damaged one is in light pink. The tip of the long β-hairpin in BHD3 (residues 599–605) is inserted into the DNA duplex and fills the gap left by the flipped-out nucleotides. The 6-4PP-linked deoxythymidine photodimer (red) is flipped out away from the protein while its deoxyadenosine partners (black) are bound by BHD2 and BHD3 of Rad4. Rad23 binds to TGD through its Rad4-binding domain (R4BD, light green). 'N' indicates the N-terminus of Rad4. 5′ and 3′ in red indicates the direction of the 6-4PP-containing DNA strand, and black the normal complementary strand. The final model contains residues 129–517 and 526–632 of Rad4 and 256–308 of Rad23. The figure was made using PyMOL Molecular Graphics System, version 2.1.1 (Schrodinger, LLC). (**E**) Electron density maps near the flipped-out 6-4PP. The Polder omit map (blue mesh) was calculated using Phenix, omitting the three DNA nucleotides containing 6-4PP including the solvent mask. Polder omit map suppresses the noise from the bulk solvent better than regular omit map (109). The polder map colored as blue mesh is contoured at 2.5σ and shown 10 Å around 6-4PP. The 2Fo-Fc map was calculated for the whole molecule, using Phenix, with density contoured at 2.0σ shown as a grey mesh, 2 Å around the DNA molecule. The positioning of the 6-4PP and its partner adenines indicate that the 6-4PP nucleotide pair is flipped out towards the major groove (block arrows). As the β-hairpin3 is inserted from the major groove, opposite of the proposed direction of 6-4PP flipping, β-hairpin3 insertion must happen after the 6-4PP flipping is at least initiated.

containing DNA duplex was generally analogous to those previously solved with (CPD-)mismatched DNA (Figure 2 and Supplementary Figure S2): Rad4 flipped out two 6-4PP damage-containing nucleotide pairs from the DNA, and the DNA in this 'open' structure was also locally bent and unwound (45,79). Such DNA binding involved all four domains (TGD, BHD1, BHD2 and BHD3) of Rad4 across its length. While the TGD and BHD1 domains encircled the fully duplexed portion of the DNA 3′ to the 6-4PP lesion, thus contacting the DNA in a damage-independent manner, the BHD2 and BHD3 domains made direct contacts with the severely distorted 6-4PP-containing DNA site. Similar to other previously solved 'open' structures, the β-hairpin3 of the BHD3 inserted into the duplex DNA and filled the gap created by the flipped out 6-4PP dinucleotide and its complementary adenines. Interestingly, the TGD, BHD1 and BHD2 domains make contacts mainly with the minor groove of the DNA while BHD3 alone faces the major groove of the 'open' DNA conformation and the long β-hairpin of BHD3 (hereafter 'β-hairpin3′) was inserted into the DNA from the major groove. Notably, 6-4PP was flipped out away from Rad4, barely making contact with the protein while the complementary adenines were bound by a narrow groove formed by the BHD2 and BHD3 interface. While this is similar to the previously solved structures, it is the first structure that shows Rad4′s flipping out and binding to purine bases on the undamaged strand, confirming that the BHD2/3 groove can indeed accept purines that are expected as partners for pyrimidine dimer lesions. The 'open' structure with 6-4PP also underscores the general mechanism of indirect recognition whereby Rad4 does not rely on a direct structural complementarity between the protein and the damaged DNA but rather indirectly senses the presence of a lesion. Such an indirect mechanism explains how a single protein complex, XPC/Rad4, could function as a common sensor for a wide variety of DNA damage repaired by NER.

The crystal structure indicates that Rad4 flips out 6-4PP nucleotide pairs towards the major rather than the minor groove

In the previously reported Rad4–DNA structures, the two flipped-out nucleotides on the 'damaged' strand were disordered leaving little trace of electron density, as they were flipped out away from Rad4 and freely exposed to the solvent. The flipped-out nucleotides in these cases were either two normal nucleotides or a CPD embedded in a 3-bp mismatch (TTT/T(CPD)). However, in this Rad4-6-4PP structure, we have observed weak yet distinct electron densities for the phosphate groups connecting the 6-4PP dinucleotides to each other and to the flanking nucleotides; We also observed electron densities for the deoxyribose of the 3′-pyrimidone (C5′, C4′ and C3′ groups), but not for the ribose of the 5′-thymidine (Figure 2E). We attribute this to the following observations. First, compared with normal dinucleotides, 6-4PP or CPD as dinucleotides are expected to be much more restricted in their movements due to the covalent linkages between the bases. In fact, the intrinsic rigidity of the 6-4PP originating from the 6-4 linkage has been previously noted (98): the conformations of the 6-4PP lesion moiety were all very similar, whether found within a duplex DNA, or as an isolated dT (6-4)dT dinucleotide, or as bound to a 6-4PP-specific antibody Fab fragment (67,98–101). We speculate that such rigidity may have contributed to the electron density observed in the current structure. Also, we note that only the 3-pyrimidone but not the 5-thymidine backbone groups engaged in van der Waals stacking with Rad4 (Arg601), which may have aided the positional stability of the 3′-base's phosphate and deoxyribose ring in the structure. Lastly, the prior structure solved with TTT/T(CPD) had not only two mismatched bases against the CPD thymine dimer but also an additional 3rd mismatch, T/T, positioned 3′ to the CPD lesion. The extra T/T mismatch showed suboptimal stacking and base pairing with relatively poor electron density and high B-factors. Therefore, it is possible that the added dynamics and instability due to the mismatches in the CPD-containing region prevented the CPD moiety (including its phosphate group) from showing observable electron density, as did the 6-4PP in a natural, matched DNA context in the present structure.

Based on the position of the positive electron density and the conformationally restricted dinucleotide structure of the 6-4PP (67), we have modeled the backbone and bases of the 6-4PP into the structure (Figure 2D, E). As mentioned earlier, the DNA duplex bound to Rad4 is shown to be locally bent and unwound at the 6-4PP lesion site and 6-4PP is bulged out of the DNA duplex. The continuity of the DNA strand containing 6-4PP and the positions of the phosphate-backbone groups necessitate that the 6-4PP moiety tightly bridges two DNA segments with severe kinks (Supplementary Figure S3). Interestingly, we also noted that the conformations of the flipped out nucleotides in the structure (particularly the two partner adenines as well as the 3′-pyrimidone in 6-4PP) suggested that the nucleotides are rotated out towards the major groove side rather than the minor groove (Figure 2E). Because the β-hairpin3 was inserted from the major groove in a direction opposite to this putative direction of the nucleotide flipping, it then suggested that the nucleotide flipping/DNA opening must be initiated before the β-hairpin3 is inserted (Figure 2E). If true, this model would complement the previous studies using the T-jump approach on mismatched DNA as a model lesion (79,80). These studies have shown that a mutant of Rad4 lacking the BHD3 domain was still capable of reaching the rate-limiting step of DNA 'opening' that leads to fully flipped-out nucleotide pairs shown in crystal structures. Another study by Van Houten and Min using atomic force microscopy and single-molecule fluorescence microscopy also showed that the Rad4 mutant lacking BHD3 can bend DNA and localize to the lesion sites (81). These results are thus congruent with the hypothesis that β-hairpin3 insertion may be a late event in the Rad4-lesion binding trajectory, which we then then closely examined through MD simulations as discussed below.

MD simulations of the initial binding between Rad4 and DNA lesions

The crystal structure, while showing how the final structure may look when Rad4 has 'recognized' the 6-4PP, begs the question of how the protein-DNA complex structure is reached as the protein and the lesion encounter each

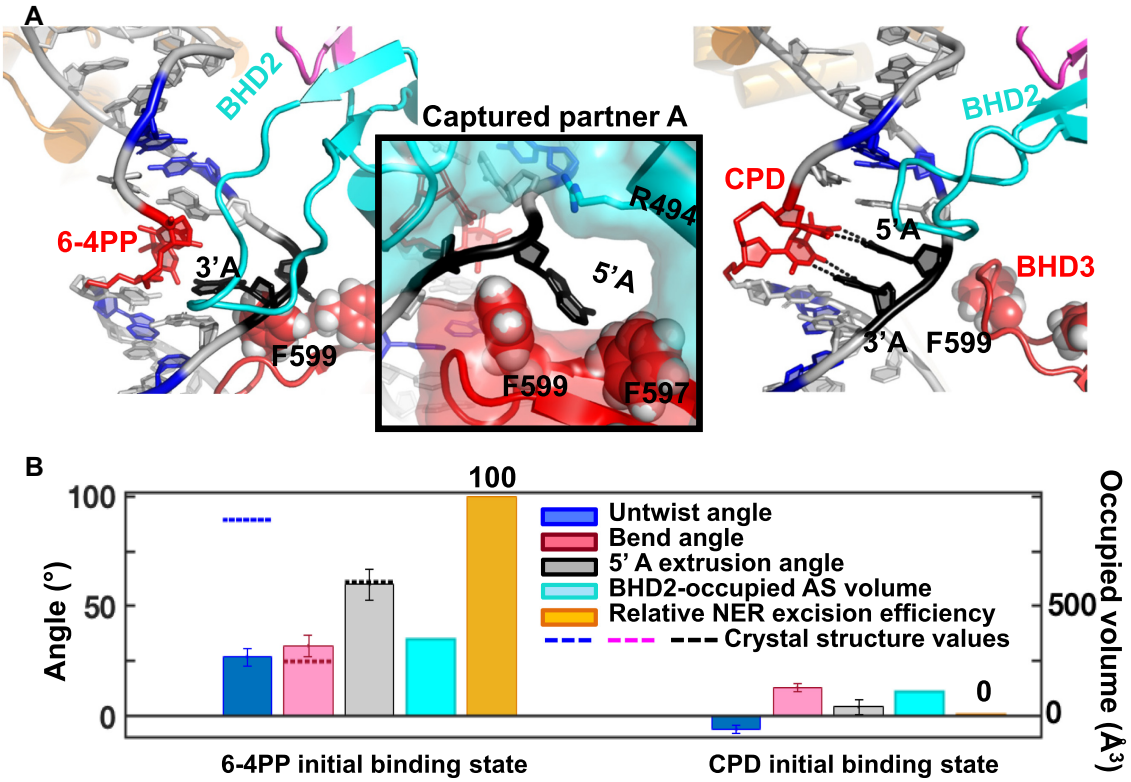


Figure 3. Rad4-UV-lesion DNA initial binding structures and their characteristics from MD simulations. (**A**) Best representative structures of the initial binding state from the MD trajectories. The structures are rendered as in Figure 2D; the end base pairs of the lesion-containing 6-mer for the calculation of untwist angles (Supplementary Figure S4B) are in blue and the side chains of F597 and F599 are in spheres. The inset depicts a zoomed-in view near the flipped out 5′dA partner opposite 4-pyrimidone (3′ T*) of 6-4PP; its binding pocket is shown as surface. R494 is shown as sticks. The black dashed lines indicate hydrogen bonds with lesion partner bases that have occupancies over 50%. (**B**) Structural characteristics for the best representative, initial binding state structures. The untwist and bend angles are shown in blue and pink, respectively, and the nucleotide flipping angles for 5′ partner A (5′dA) is in gray. The corresponding values for the 6-4PP-bound to Rad4 in the crystal structure are indicated as dashed lines. The BHD2-occupied alpha space (AS) volumes and relative NER excision efficiencies for the 6-4PP and CPD-containing duplexes are in cyan and orange. The NER excision efficiency with 6-4PP was assigned a relative value of 100 (110). The error bars indicate the standard deviations for the block average values of the measured angles (see SI Methods)

other in solution. In fact, the process by which this structure is reached must hold the key as to whether or not the final structure can be reached in the first place. Importantly, recent MD simulation studies by Mu *et al.* from the Broyde group have shown remarkable connections between the early stage interactions between Rad4 and lesion-containing DNA duplexes and the final recognition and repair efficiencies (82,102). Motivated by these studies and to obtain mechanistic insights into the UV-lesion recognition process, we performed 4-μs MD simulations for the initial binding of Rad4 to a 6-4PP-containing DNA duplex as well as to a CPD-containing duplex. The 'docking complexes' used as the starting models were generated based on a combination of the lesion-bound Rad4–Rad23, the free Rad4–Rad23 and the free lesion-containing DNA structures as previously described (SI Methods, Supplementary Figure S4A) (82,102,103). The thymine-derived photodimer lesions were paired with the normal partner adenines and were placed within the same DNA sequence context as in the crystal structure (Supplementary Figure S4B). The 'docking complex' represents the state of Rad4 and DNA prior to initial engagement of the BHD2 and BHD3 domains. The resulting MD simulations for the binding of 6-4PP and CPD are shown in Movies S1 and S2, respectively. To map out distinct conformational ensembles presented along the simulation trajectories, we carried out principal component analysis (PCA), which then defined 6 distinct structural clusters for the 6-4PP and four clusters for the CPD trajectory (Supplementary Figure S5). Each cluster is an ensemble of structures that exhibits similar structural dynamics and could represent a sub-stage upon Rad4 initial binding to the lesion-containing duplex. To gain insights into the lesion recognition process in detail, we have also computed various structural parameters around the lesion sites over the entire trajectories (Supplementary Figure S6–S11), which are discussed further below.

6-4PP, but not CPD, undergoes remarkable DNA untwisting, bending and nucleotide-flipping towards the major groove

First, we compare the conformational clusters which stably dominated the trajectories between 3 and 4 μs for 6-4PP (Figure 3A, Supplementary Figure S5A blue (g5) & Movie S3) and for CPD-binding (Figure 3A, Supplementary Figure S5B green (g3) & Movie S4). These clusters, which we refer to as 'initial binding state', best illustrate the key dif-

ferences between 6-4PP and CPD in their initial binding to Rad4 to give us insights into the differences in their recognition by Rad4 and repair in NER.

DNA untwisting. The 6-4PP-DNA in the crystal structure showed DNA untwisting around the lesion. Prior experimental (80) and computational studies (102,104) have also shown that Rad4 untwists/unwinds DNA as it probes DNA nonspecifically for the presence of a lesion in the initial binding state. To examine whether different 'untwisting' may contribute to the varying Rad4-binding specificities for the UV lesions, we calculated the *un*twist angles along the trajectory of Rad4 initial binding to the lesion-containing duplexes: Untwist = Twist $_{\text{pre-BHD2 engagement}}$ –Twist, as illustrated in Supplementary Figure S6. Twist$_{\text{pre-BHD2 engagement}}$ is the ensemble average twist angle of the lesion-containing 6-mer during the first 1 ns of production MD during which there is no significant protein binding-induced changes. Positive values indicate untwisting and negative values indicate over-twisting. Our analyses show that the DNA is untwisted around the 6-4PP by 27 ± 4° during the initial binding with Rad4, which was smaller than the untwist angle of 89° shown in the crystal structure but was significantly larger than that observed with CPD: the CPD-containing DNA shows no untwisting but rather a slight over-twisting of 6 ± 2° (Figure 3B).

DNA bending. We also observed sharp differences between 6-4PP and CPD in their DNA bend angles around the lesion sites throughout the MD trajectories (see Supplementary Methods, Figure 3B and Supplementary Figure S6). The 6-4PP-containing DNA exhibited significant bending of 32 ± 5° in the initial binding state, with the sequence 5′ to the lesion bent *from* the minor groove side, accompanied by extensive engagement of the Rad4 BHD2 domain with the minor groove of DNA (Figure 3 & Movie S3). The bend angle is larger than those of the DNA in the 'docking' model (23°) or in the present crystal structure (25°). On the other hand, the CPD-containing duplex was bent by only 13 ± 2° with the sequence 5′ to the lesion bent in a direction opposite to that in the 6-4PP structure, disallowing the lesion site to be engaged with Rad4 in the 'open' conformation (Figure 3, Supplementary Figure S6 & Movie S4). The bending of 6-4PP DNA was also highly dynamic compared to that of CPD during BHD2 binding, which was limited for CPD (described below). Furthermore, the 6-4PP bend angle correlated positively with the degree of untwisting (correlation coefficient $r = 0.56$) for the ensemble from 1 to 4 μs, indicating potential coupling of the two motions (Supplementary Figure S6).

Nucleotide-flipping. Perhaps the most dramatic difference between 6-4PP and CPD during the initial binding to Rad4 was that only 6-4PP-containing nucleotide pairs were extruded from the DNA duplex and lost most of their Watson-Crick hydrogen-bond pairing, while CPD did not (Movie S1 & S2, Figure 3 and Supplementary Figure S7). Furthermore, the 6-4PP:AA nucleotide pairs all extrude towards the major groove, as indicated by the directions of the nucleotide-flipping pseudo-dihedral angle changes (Figure 3 and Supplementary Figure S7B) (105). This directly supports our inference from the current crystal structure of the Rad4–6-4PP complex. During the extrusion processes, 6-4PP extruded first followed by the 5′ partner dA (5′dA) which flipped out towards the Rad4 protein. Finally, the 3′ partner dA (3′dA) partially extruded episodically while occasionally maintaining one hydrogen bond with the 6-4PP (Figure 3 and Supplementary Figures S7 and S8). In the major groove, the 6-4PP and its partner nucleotides manifested dynamics, and the 5′dA occasionally rotates around the glycosidic bond from the *anti* to the *syn* conformation (Supplementary Figure S9), as adopted in the crystal structures. By contrast, the CPD lesion and partner bases did not extrude and steadily maintained their Watson-Crick base pairing as well as the normal, *anti*-glycosidic bond (Figure 3, Supplementary Figures S7–S9).

Altogether, these analyses demonstrated that the DNA containing 6-4PP, but not CPD, becomes untwisted and bent and has its damage-containing nucleotide pairs flipped out upon its initial binding to Rad4. Next, we describe the key Rad4–DNA interactions that underlie the observed conformational changes of the 6-4PP DNA.

6-4PP:AA nucleotide pair extrusion involve significant engagement of the Rad4 BHD2 domain and manifest compensatory exchanges of van der Waals interactions with interacting partners in DNA and Rad4

Previous MD studies indicated that lesion recognition correlated positively with how well BHD2 initially engaged with the lesion site from the minor groove side of the DNA (104). To examine this for 6-4PP and CPD, we calculated the alpha space (AS) volume occupied by BHD2 in the minor groove of DNA (Supplementary Figure S10): the AS volume reflects the curvature and surface area of the DNA minor groove bound by BHD2 (106). The AS volume for the 6-4PP-binding was 349 (Å^3) in the initial binding state and was consistently high from early in the binding trajectory, but it was only 110 (Å^3) for the CPD-binding (Figure 3 and Supplementary Figure S10). Thus, the Rad4 BHD2 interacted more extensively with the minor groove around 6-4PP than around CPD. This binding was achieved by multiple hydrogen bonding and hydrophobic interactions between the 6-4PP-containing DNA and BHD2 (particularly involving Val517) while the CPD manifested only limited interactions (Supplementary Figure S11, Supplementary Table S3, Movies S3 and S4). Notably, the analyses of the *van der Waals* interactions involving the lesion, its partner bases, their neighboring nucleotides as well as BHD2-BHD3 show that there is successive handover between the interacting partners where the loss of one interaction is compensated by a new interaction with another party. Furthermore, the compensatory exchanges were carried out in a manner that directly facilitate nucleotide extrusions and partner base flipping (Supplementary Figure S12, see SI Discussion).

Visualizing the full lesion recognition trajectory

Combining the crystallographic and MD structures together, we can start constructing a plausible high-resolution

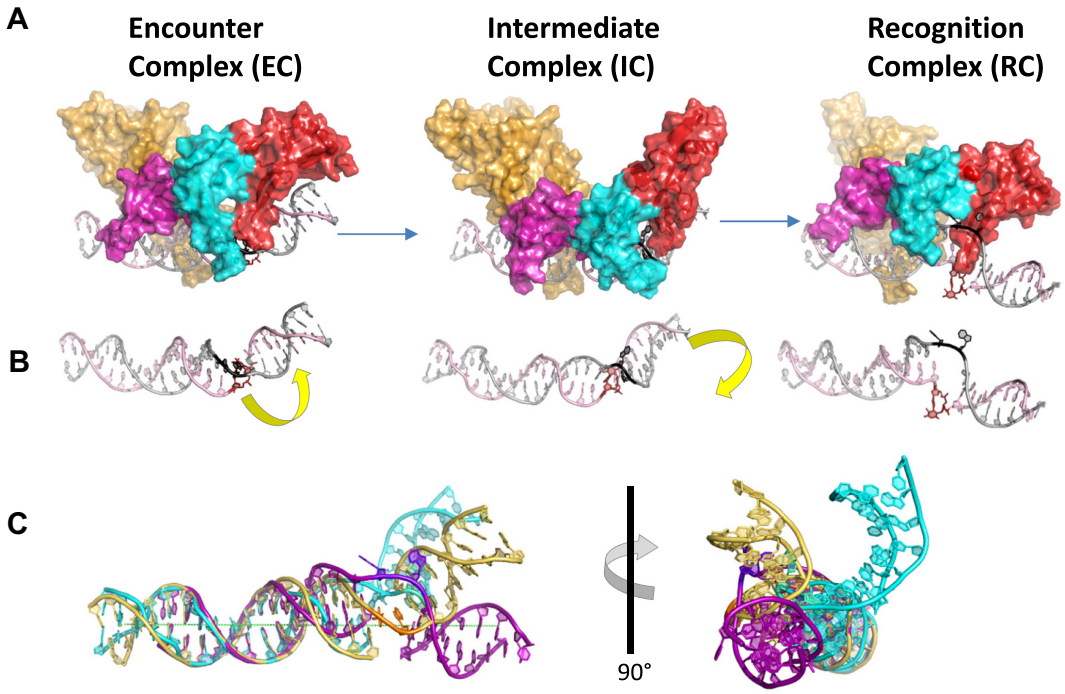


Figure 4. Structural models for the 6-4PP lesion recognition trajectory by Rad4. The EC, IC (g5) and RC structures were superposed over the undamaged DNA duplex regions bound by the TGD domain of Rad4 (residues 3–11 in Figure 2C) and each structure is shown as a protein–DNA complex (**A**) or DNA only (**B**) from the same orientation. (**C**) The superposed DNA structures are shown in two orientations. EC is in yellow, IC in cyan and RC in purple. The green dotted line indicates the helical axis of the DNA extending from the undamaged B-DNA duplex region.

3D binding trajectory for Rad4 and 6-4PP. While the Rad4–6-4PP crystal structure serves as the final complex formed ('recognition complex', RC), the docking model for MD simulation represents the initial encounter between the protein and the lesion (encounter complex, EC) where the DNA structures resemble the free DNA structure the most. The MD structural clusters that appear after the initial equilibration can be considered as early intermediate complexes (ICs) between EC and RC along the recognition trajectory. Figure 4 and Movie S5 depict the structural progression from EC to IC (taken from the g5 in Supplementary Figure S5A) and then to the RC. Figure 4B depicts the DNA duplexes extracted from the superposed structures. As expected, all three DNAs in these complexes show bent conformations but the directions were all distinct from one another. Notably, IC showed most bending among the three structures owing to the engagement of BHD2 from the minor groove side, which bends the DNA towards the major groove side. The RC, on the other hand, was relatively less bent but most unwound, accompanied by the flipping out ('opening') of both of the 6-4PP-containing nucleotide pairs that is stabilized by the BHD3 β-hairpin insertion. The IC structure(s) clearly indicate that Rad4 uses DNA bending and untwisting as passage to achieve the RC's 'open' structure.

Altogether, this study enables us to visualize the structural trajectory for Rad4′s photolesion binding in unprecedented detail and provides key structural insights into the remarkable differences between the 6-4PP and CPD recognition and repair by NER.

DISCUSSION

Indirect recognition of 6-4PP by Rad4 is distinct from the recognition of 6-4PP by other 6-4PP-binding proteins

A limited number of 3D structures have previously been determined for other 6-4PP-binding proteins bound to 6-4PP-containing double-stranded DNA, including UV-DDB (28), 6-4 photolyase (107) and anti-6-4PP antibodies (108) using X-ray crystallography (reviewed in (98)). In these complex structures, the protein-bound DNA duplexes were all kinked around the 6-4PP by ∼40–90° with the 6-4PP flipped out towards a cavity or an active site of the protein which makes direct contacts with 6-4PP. The direct contacts between these proteins and 6-4PP would be pertinent for the functions of the proteins, which entail specific recognition of the 6-4PP (e.g. antibodies raised against 6-4PP or the 6-4PP photolyase that carries out the photo-reversal of 6-4PP). In the Rad4–6-4PP structure, we observed that 6-4PP was flipped out from the DNA duplex similarly as in other structures, but in contrast to others, Rad4 did not make direct contacts with the lesion. Such a unique binding mode compels the protein to rely on an *indirect recognition* strategy and allows the protein to bind to an extraordinarily broad range of DNA damage (45,77,79).

MD trajectories show sequential extrusion of both nucleotides complementary to 6-4PP towards the major groove upon the initial binding with Rad4

Previous MD simulation studies by Mu *et al.* from the Broyde group have examined the initial Rad4–DNA inter-

actions for various DNA lesions derived from polycyclic aromatic chemicals (102–104). The studies revealed the value of the computational approach in uncovering the molecular mechanisms of Rad4 lesion recognition. In particular, it revealed that the NER rates of aromatic adduct lesions had strong positive correlations with the degrees of DNA untwisting and of BHD2-minor groove engagement in the initial binding states reached as early as 300 ns and stable up to 1.5 μs of MD simulations (81). In the present study, we extended this approach to the Rad4-binding with 6-4PP and CPD (all in a natural, matched DNA context) and carried out simulations up to 4 μs. This report is the first that examined the initial binding state MD trajectories of small intra-strand crosslink dimeric lesions, which do not entail a large aromatic modification to nucleobases. While we observed a common pattern of binding characteristics that contrast sharply between well-repaired and repair-resistant lesions, the current study also adds unique insights for 6-4PP recognition, derived from the particular 6-4PP structure. Specifically, here we observe the flipping of the 5′ partner dA binding into a BHD2/BHD3 groove at ∼2 μs, followed by episodic partial extrusion of the 3′ partner dA at ∼2.5 μs to approach the hairpin tip of BHD3. On the other hand, just one nucleotide had been observed to flip in the previous 1.5 μs simulations for polycyclic aromatic moieties (82,102). The distorted original hydrogen bonding of 6-4PP with dA partners promoted the extrusion process particularly for the 5′ partner dA. The unstacking and full flipping of the 5′ dA leads the 3′ dA to partially and episodically extrude, while occasionally maintaining one hydrogen bond. 6-4PP itself also showed dynamical partial extrusions. Remarkably, in all these trajectories, the nucleotide extrusion or 'flipping-out' progressed towards the major groove, as hinted from the Rad4–6-4PP crystal structure which represents the final state for a Rad4-recognized lesion. Once both partner adenines are fully flipped out, they would both be captured by a binding pocket between BHD2 and BHD3 in Rad4 prior to the insertion of the β-hairpin3 into the DNA duplex shown in the crystal structure and in our previous full pathway simulations (102,103). The MD analyses thus further solidify the model that the major groove extrusion of both nucleotide pairs may begin before the β-hairpin3 insertion (80,103). Interestingly, the Rad4-binding-induced nucleotide flipping towards the major groove was also observed with the well-repaired 14*R*-(+)-*trans-anti*-dibenzo[*a,l*]-pyrene-N^2-dG (14*R*-DB[*a,l*]P-dG) lesion paired with normal dC (82). The general observation of partner nucleotide capture has been described for a number of well-repaired lesions: 10*R*-(+)-*cis-anti*-benzo[*a*]pyrene-N^2-dG paired with normal dC (*cis*-B[*a*]P-dG:dC), N-(deoxyguanosin-8-yl)-2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine paired with normal dC (PhIP-C8-dG:dC) and 14*R*-DB[*a,l*]P-dG:dC, and may well be a characteristic of lesions with high NER efficiencies (96).

New in this study, we have also found that the untwisting angles are positively correlated with the DNA bend angles particularly after 1 μs, indicating that these motions in the initial binding states exhibit mechanistic coupling. This coupling may be possible as BHD2 engages from the minor groove side of the DNA while BHD3 faces the major groove, encouraging bending directed toward the open complex upon untwisting and vice versa.

In contrast to the well-recognized/repaired 6-4PP, the unrecognizable CPD resisted unwinding by the Rad4′s BHD2 and did not engage much with BHD2. These features are congruent with the features shown by other repair-resistant lesions such as *cis*-B[a]P-dG missing its partner dC or mispaired with a dA or 14*R*-DB[*a,l*]P-dA that was paired normally with a dT (82). Furthermore, CPD showed that all of its Watson-Crick pairing for both base pairs was maintained throughout the simulation as was also the case for the repair-resistant 14*R*-DB[*a,l*]P-dA:dT (82). Some bending of ∼ 27° in the CPD DNA was observed before 1.5 μs but the DNA maintained mostly straight forms afterwards, consistent with the lack of untwisting. It is intriguing to speculate that CPD may avoid recognition by resisting the coupled untwisting/bending, for instance, by being able to bend without the untwisting that is needed for nucleotide extrusion.

Implications for the 'kinetic gating' mechanism

The current study also corroborates the 'kinetic gating' mechanism that we have previously proposed (79). In the 'kinetic gating' model, we proposed that Rad4-induced DNA duplex 'opening' must happen before the protein translocates away from the DNA site in order for a DNA site to be recognized. The kinetic studies using T-jump spectroscopy had revealed that the nonspecific interrogation involving untwisting motion in the DNA is in the order of 100–500 μs whereas the rate-determining full 'opening' step was ∼5–10 ms for mismatched DNA model lesions (79,80). Studies using single molecule microscopy revealed that the residence time of the protein during nonspecific, 1D diffusional search is 1–600 μs per DNA base pair (81). It is currently not feasible to simulate time scales in the 100 μs range with all-atom MD simulations and the computationally demanding but most accurate explicit solvation employed here; furthermore, it remains to be determined how the time scales in the MD trajectory exactly translate to the experimentally observed ones. We also suggest that these time scales will differ depending on the DNA lesions and sequence contexts. However, the 3D movies of molecular motions obtained through the MD studies here already successfully manifest how 6-4PP would be kinetically more likely to be 'opened' before Rad4 diffuses away than CPD and provides a firm foundation for future studies aimed at directly bridging the computational and experimental studies.

Implication for the mechanism of lesion recognition by UV-DDB

The present study shows that the BHD2 engagement with the DNA minor groove around the lesion site is a critical component leading to the lesion-specific binding. Interestingly, the UV-DDB complex is also shown to engage with the DNA minor groove and pushes out the nucleotides towards the major groove. The space previously occupied by the lesion is filled by a three-residue plug (Phe371, Gln372, His373) inserted from the minor groove side (28). The nucleotide pairs such as CPD:AA are being partially flipped

out and this leaves the partner purines against the photolesions available for capture by XPC, as the DDB2 protein is facing the damaged strand, partially enclosing the damaged photodimers. Interestingly, when the DDB2–DNA structure is superposed to the initial binding state of CPD with Rad4, the DNA bending directions are consistent with each other, with the DDB2 and Rad4 facing opposite sides of the DNA (Supplementary Figure S13). BHD2 also overlaps with DDB2′s plug residues, which indicates that the lesion hand-over from UV-DDB to XPC may involve replacing DDB2 with the BHD2 of XPC at some point during the process (33).

In sum, our study provides new insights into the lesion recognition process by XPC/Rad4 and NER. We envision that the structural and mechanistic principles gathered from this study will be applicable to a broad range of protein–nucleic acid binding and recognition processes that occur in cells in the intricate chromatin context and will illuminate studies geared towards developing strategies to modulate these interactions for clinical interventions.

DATA AVAILABILITY

The structural coordinates have been deposited with PDB code 6CFI.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr. Nikola P. Pavletich for initiating this project and the members of the Min and Broyde groups for their support. We are grateful to Ms. Chunyan (Alice) Cao for help with synthesis of the thymidine dimer methyl ester and HPLC purifications and to Dr. George Sukenick for help with NMR.

FUNDING

National Science Foundation (NSF) [MCB-1412692 to J.-H.M.]; National Institutes of Health [R21-ES028384 to J.-H.M. and R01-ES025987 to S.B.]; This work used the Extreme Science and Engineering Discovery Environment (XSEDE), supported by NSF [MCB-060037 to S.B.] and high performance computing resources of New York University (NYU-ITS); The Organic synthesis Core at MSKCC is partially supported by an NCI grant [P30 CA008748]. Funding for open access charge: NSF/NIH grants.
Conflict of interest statement. None declared.

REFERENCES

1. Friedberg,E.C., Walker,G.C., Siede,W., Wood,R.D., Schultz,R.A. and Ellenberger,T. (2006) *DNA Repair and Mutagenesis*. 2nd edn. ASM Press, Washington, D.C.
2. Ganesan,A. and Hanawalt,P. (2016) Photobiological origins of the field of genomic maintenance. *Photochem. Photobiol.*, **92**, 52–60.
3. Mao,P., Wyrick,J.J., Roberts,S.A. and Smerdon,M.J. (2017) UV-Induced DNA damage and mutagenesis in chromatin. *Photochem. Photobiol.*, **93**, 216–228.
4. Cadet,J. and Douki,T. (2018) Formation of UV-induced DNA damage contributing to skin cancer development. *Photochem. Photobiol. Sci.*, **17**, 1816–1841.
5. Jackson,S.P. and Bartek,J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071–1078.
6. Scharer,O.D. (2013) Nucleotide excision repair in eukaryotes. *Cold Spring Harb. Perspect. Biol.*, **5**, a012609.
7. Sugasawa,K. (2016) Molecular mechanisms of DNA damage recognition for mammalian nucleotide excision repair. *DNA Repair (Amst.)*, **44**, 110–117.
8. Puumalainen,M.R., Ruthemann,P., Min,J.H. and Naegeli,H. (2016) Xeroderma pigmentosum group C sensor: unprecedented recognition strategy and tight spatiotemporal regulation. *Cell. Mol. Life Sci.*, **73**, 547–566.
9. Gillet,L.C. and Scharer,O.D. (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem. Rev.*, **106**, 253–276.
10. Besaratinia,A., Yoon,J.I., Schroeder,C., Bradforth,S.E., Cockburn,M. and Pfeifer,G.P. (2011) Wavelength dependence of ultraviolet radiation-induced DNA damage as determined by laser irradiation suggests that cyclobutane pyrimidine dimers are the principal DNA lesions produced by terrestrial sunlight. *FASEB J.*, **25**, 3079–3091.
11. You,Y.H., Lee,D.H., Yoon,J.H., Nakajima,S., Yasui,A. and Pfeifer,G.P. (2001) Cyclobutane pyrimidine dimers are responsible for the vast majority of mutations induced by UVB irradiation in mammalian cells. *J. Biol. Chem.*, **276**, 44688–44694.
12. Protic-Sabljic,M., Tuteja,N., Munson,P.J., Hauser,J., Kraemer,K.H. and Dixon,K. (1986) UV light-induced cyclobutane pyrimidine dimers are mutagenic in mammalian cells. *Mol. Cell. Biol.*, **6**, 3349–3356.
13. Hu,J. and Adar,S. (2017) The cartography of UV-induced DNA damage formation and DNA repair. *Photochem. Photobiol.*, **93**, 199–206.
14. Karentz,D. (2015) Beyond xeroderma pigmentosum: DNA damage and repair in an ecological context. A tribute to James E. Cleaver. *Photochem. Photobiol.*, **91**, 460–474.
15. Cleaver,J.E., Lam,E.T. and Revet,I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat. Rev. Genet.*, **10**, 756–768.
16. Kraemer,K.H., Patronas,N.J., Schiffmann,R., Brooks,B.P., Tamura,D. and DiGiovanna,J.J. (2007) Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship. *Neuroscience*, **145**, 1388–1396.
17. Spivak,G. (2015) Nucleotide excision repair in humans. *DNA Repair (Amst.)*, **36**, 13–18.
18. Brueckner,F., Hennecke,U., Carell,T. and Cramer,P. (2007) CPD damage recognition by transcribing RNA polymerase II. *Science*, **315**, 859–862.
19. Xu,J., Lahiri,I., Wang,W., Wier,A., Cianfrocco,M.A., Chong,J., Hare,A.A., Dervan,P.B., DiMaio,F., Leschziner,A.E. *et al.* (2017) Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature*, **551**, 653–657.
20. Sanz-Murillo,M., Xu,J., Belogurov,G.A., Calvo,O., Gil-Carton,D., Moreno-Morcillo,M., Wang,D. and Fernandez-Tornero,C. (2018) Structural basis of RNA polymerase I stalling at UV light-induced DNA damage. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, 8972–8977.
21. Li,W., Selvam,K., Ko,T. and Li,S. (2014) Transcription bypass of DNA lesions enhances cell survival but attenuates transcription coupled DNA repair. *Nucleic Acids Res.*, **42**, 13242–13253.
22. van Hoffen,A., Venema,J., Meschini,R., van Zeeland,A.A. and Mullenders,L.H. (1995) Transcription-coupled repair removes both cyclobutane pyrimidine dimers and 6-4 photoproducts with equal efficiency and in a sequential way from transcribed DNA in xeroderma pigmentosum group C fibroblasts. *EMBO J.*, **14**, 360–367.
23. Adar,S., Hu,J., Lieb,J.D. and Sancar,A. (2016) Genome-wide kinetics of DNA excision repair in relation to chromatin state and mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E2124–E2133.
24. Friedberg,E.C. (2001) How nucleotide excision repair protects against cancer. *Nat. Rev. Cancer*, **1**, 22–33.
25. Wood,R.D. (1999) DNA damage recognition during nucleotide excision repair in mammalian cells. *Biochimie*, **81**, 39–44.

26. Sugasawa,K., Ng,J.M., Masutani,C., Iwai,S., van der Spek,P.J., Eker,A.P., Hanaoka,F., Bootsma,D. and Hoeijmakers,J.H. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol. Cell*, **2**, 223–232.
27. Nishi,R., Okuda,Y., Watanabe,E., Mori,T., Iwai,S., Masutani,C., Sugasawa,K. and Hanaoka,F. (2005) Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Mol. Cell. Biol.*, **25**, 5664–5674.
28. Scrima,A., Konickova,R., Czyzewski,B.K., Kawasaki,Y., Jeffrey,P.D., Groisman,R., Nakatani,Y., Iwai,S., Pavletich,N.P. and Thoma,N.H. (2008) Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell*, **135**, 1213–1223.
29. Ghodke,H., Wang,H., Hsieh,C.L., Woldemeskel,S., Watkins,S.C., Rapic-Otrin,V. and Van Houten,B. (2014) Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E1862–E1871.
30. Yeh,J.I., Levine,A.S., Du,S., Chinte,U., Ghodke,H., Wang,H., Shi,H., Hsieh,C.L., Conway,J.F., Van Houten,B. *et al.* (2012) Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E2737–E2746.
31. 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.
32. Sugasawa,K. (2011) Multiple DNA damage recognition factors involved in mammalian nucleotide excision repair. *Biochemistry (Mosc.)*, **76**, 16–23.
33. Sugasawa,K., Okuda,Y., Saijo,M., Nishi,R., Matsuda,N., Chu,G., Mori,T., Iwai,S., Tanaka,K. and Hanaoka,F. (2005) UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell*, **121**, 387–400.
34. Fitch,M.E., Nakajima,S., Yasui,A. and Ford,J.M. (2003) In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *J. Biol. Chem.*, **278**, 46906–46910.
35. Batty,D., Rapic'-Otrin,V., Levine,A.S. and Wood,R.D. (2000) Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *J. Mol. Biol.*, **300**, 275–290.
36. Sugasawa,K., Okamoto,T., Shimizu,Y., Masutani,C., Iwai,S. and Hanaoka,F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.*, **15**, 507–521.
37. Hey,T., Lipps,G., Sugasawa,K., Iwai,S., Hanaoka,F. and Krauss,G. (2002) The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay. *Biochemistry*, **41**, 6583–6587.
38. Young,A.R., Chadwick,C.A., Harrison,G.I., Hawk,J.L., Nikaido,O. and Potten,C.S. (1996) The in situ repair kinetics of epidermal thymine dimers and 6-4 photoproducts in human skin types I and II. *J. Invest. Dermatol.*, **106**, 1307–1313.
39. Canturk,F., Karaman,M., Selby,C.P., Kemp,M.G., Kulaksiz-Erkmen,G., Hu,J., Li,W., Lindsey-Boltz,L.A. and Sancar,A. (2016) Nucleotide excision repair by dual incisions in plants. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 4706–4710.
40. Uchida,A., Sugasawa,K., Masutani,C., Dohmae,N., Araki,M., Yokoi,M., Ohkuma,Y. and Hanaoka,F. (2002) The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH. *DNA Repair (Amst.)*, **1**, 449–461.
41. Yokoi,M., Masutani,C., Maekawa,T., Sugasawa,K., Ohkuma,Y. and Hanaoka,F. (2000) The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. *J. Biol. Chem.*, **275**, 9870–9875.
42. Kusakabe,M., Onishi,Y., Tada,H., Kurihara,F., Kusao,K., Furukawa,M., Iwai,S., Yokoi,M., Sakai,W. and Sugasawa,K. (2019) Mechanism and regulation of DNA damage recognition in nucleotide excision repair. *Genes Environ.*, **41**, 2.
43. Li,C.L., Golebiowski,F.M., Onishi,Y., Samara,N.L., Sugasawa,K. and Yang,W. (2015) Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Mol. Cell*, **59**, 1025–1034.
44. Riedl,T., Hanaoka,F. and Egly,J.M. (2003) The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.*, **22**, 5293–5303.
45. Min,J.H. and Pavletich,N.P. (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature*, **449**, 570–575.
46. Volker,M., Mone,M.J., Karmakar,P., van Hoffen,A., Schul,W., Vermeulen,W., Hoeijmakers,J.H., van Driel,R., van Zeeland,A.A. and Mullenders,L.H. (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Mol. Cell*, **8**, 213–224.
47. Luijsterburg,M.S., von Bornstaedt,G., Gourdin,A.M., Politi,A.Z., Mone,M.J., Warmerdam,D.O., Goedhart,J., Vermeulen,W., van Driel,R. and Hofer,T. (2010) Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. *J. Cell Biol.*, **189**, 445–463.
48. Sugasawa,K., Akagi,J., Nishi,R., Iwai,S. and Hanaoka,F. (2009) Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. *Mol. Cell*, **36**, 642–653.
49. Sugasawa,K., Shimizu,Y., Iwai,S. and Hanaoka,F. (2002) A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA Repair (Amst.)*, **1**, 95–107.
50. Bootsma,D., Kraemer,K.H., Cleaver,J. and Hoeijmakers,J.H. (1998) In: Vogelstein B, and Kinzler,KW (eds). *The Genetic Basis of Human Cancer*. McGraw-Hill, NY, pp. 245–274.
51. Chavanne,F., Broughton,B.C., Pietra,D., Nardo,T., Browitt,A., Lehmann,A.R. and Stefanini,M. (2000) Mutations in the XPC gene in families with xeroderma pigmentosum and consequences at the cell, protein, and transcript levels. *Cancer Res.*, **60**, 1974–1982.
52. Perera,D., Poulos,R.C., Shah,A., Beck,D., Pimanda,J.E. and Wong,J.W. (2016) Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. *Nature*, **532**, 259–263.
53. Budden,T., Davey,R.J., Vilain,R.E., Ashton,K.A., Braye,S.G., Beveridge,N.J. and Bowden,N.A. (2016) Repair of UVB-induced DNA damage is reduced in melanoma due to low XPC and global genome repair. *Oncotarget*, **7**, 60940–60953.
54. Murray,H.C., Maltby,V.E., Smith,D.W. and Bowden,N.A. (2015) Nucleotide excision repair deficiency in melanoma in response to UVA. *Exp. Hematol. Oncol.*, **5**, 6.
55. Dong,T.K., Ona,K., Scandurra,A.E., Demetriou,S.K. and Oh,D.H. (2016) Deficient nucleotide excision repair in squamous Cell carcinoma cells. *Photochem. Photobiol.*, **92**, 760–766.
56. Kusumoto,R., Masutani,C., Sugasawa,K., Iwai,S., Araki,M., Uchida,A., Mizukoshi,T. and Hanaoka,F. (2001) Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro. *Mutat. Res.*, **485**, 219–227.
57. Reardon,J.T., Mu,D. and Sancar,A. (1996) Overproduction, purification, and characterization of the XPC subunit of the human DNA repair excision nuclease. *J. Biol. Chem.*, **271**, 19451–19456.
58. Bunick,C.G., Miller,M.R., Fuller,B.E., Fanning,E. and Chazin,W.J. (2006) Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein. *Biochemistry*, **45**, 14965–14979.
59. Trego,K.S. and Turchi,J.J. (2006) Pre-steady-state binding of damaged DNA by XPC-hHR23B reveals a kinetic mechanism for damage discrimination. *Biochemistry*, **45**, 1961–1969.
60. Cai,Y., Patel,D.J., Geacintov,N.E. and Broyde,S. (2009) Differential nucleotide excision repair susceptibility of bulky DNA adducts in different sequence contexts: hierarchies of recognition signals. *J. Mol. Biol.*, **385**, 30–44.
61. Kropachev,K., Kolbanovskii,M., Cai,Y., Rodriguez,F., Kolbanovskii,A., Liu,Y., Zhang,L., Amin,S., Patel,D., Broyde,S. *et al.* (2009) The sequence dependence of human nucleotide excision repair efficiencies of benzo[a]pyrene-derived DNA lesions: insights into the structural factors that favor dual incisions. *J. Mol. Biol.*, **386**, 1193–1203.
62. Cai,Y., Kropachev,K., Xu,R., Tang,Y., Kolbanovskii,M., Kolbanovskii,A., Amin,S., Patel,D.J., Broyde,S. and Geacintov,N.E. (2010) Distant neighbor base sequence context effects in human nucleotide excision repair of a benzo[a]pyrene-derived DNA lesion. *J. Mol. Biol.*, **399**, 397–409.
63. Mu,H., Kropachev,K., Wang,L., Zhang,L., Kolbanovskiy,A., Kolbanovskiy,M., Geacintov,N.E. and Broyde,S. (2012) Nucleotide

excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies. *Nucleic Acids Res.*, **40**, 9675–9690.
64. Ding,S., Kropachev,K., Cai,Y., Kolbanovskiy,M., Durandina,S.A., Liu,Z., Shafirovich,V., Broyde,S. and Geacintov,N.E. (2012) Structural, energetic and dynamic properties of guanine(C8)-thymine(N3) cross-links in DNA provide insights on susceptibility to nucleotide excision repair. *Nucleic Acids Res.*, **40**, 2506–2517.
65. Cai,Y., Patel,D.J., Geacintov,N.E. and Broyde,S. (2007) Dynamics of a benzo[a]pyrene-derived guanine DNA lesion in TGT and CGC sequence contexts: enhanced mobility in TGT explains conformational heterogeneity, flexible bending, and greater susceptibility to nucleotide excision repair. *J. Mol. Biol.*, **374**, 292–305.
66. Kim,J.K. and Choi,B.S. (1995) The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidylyl(3′→5′)thymidine by NMR and relaxation matrix refinement. *Eur. J. Biochem.*, **228**, 849–854.
67. Kim,J.K., Patel,D. and Choi,B.S. (1995) Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity. *Photochem. Photobiol.*, **62**, 44–50.
68. Dehez,F., Gattuso,H., Bignon,E., Morell,C., Dumont,E. and Monari,A. (2017) Conformational polymorphism or structural invariance in DNA photoinduced lesions: implications for repair rates. *Nucleic Acids Res.*, **45**, 3654–3662.
69. Kemmink,J., Boelens,R., Koning,T.M., Kaptein,R., van der Marel,G.A. and van Boom,J.H. (1987) Conformational changes in the oligonucleotide duplex d(GCGTTGCG) x d(CGCAACGC) induced by formation of a cis-syn thymine dimer. A two-dimensional NMR study. *Eur. J. Biochem.*, **162**, 37–43.
70. McAteer,K., Jing,Y., Kao,J., Taylor,J.S. and Kennedy,M.A. (1998) Solution-state structure of a DNA dodecamer duplex containing a Cis-syn thymine cyclobutane dimer, the major UV photoproduct of DNA. *J. Mol. Biol.*, **282**, 1013–1032.
71. Wang,C.I. and Taylor,J.S. (1991) Site-specific effect of thymine dimer formation on dAn.dTn tract bending and its biological implications. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 9072–9076.
72. Park,H., Zhang,K., Ren,Y., Nadji,S., Sinha,N., Taylor,J.S. and Kang,C. (2002) Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 15965–15970.
73. Krasikova,Y.S., Rechkunova,N.I., Maltseva,E.A., Anarbaev,R.O., Pestryakov,P.E., Sugasawa,K., Min,J.H. and Lavrik,O.I. (2013) Human and yeast DNA damage recognition complexes bind with high affinity DNA structures mimicking in size transcription bubble. *J. Mol. Recognit.*, **26**, 653–661.
74. Zhang,E.T., He,Y., Grob,P., Fong,Y.W., Nogales,E. and Tjian,R. (2015) Architecture of the human XPC DNA repair and stem cell coactivator complex. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 14817–14822.
75. Gunz,D., Hess,M.T. and Naegeli,H. (1996) Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *J. Biol. Chem.*, **271**, 25089–25098.
76. Geacintov,N.E., Broyde,S., Buterin,T., Naegeli,H., Wu,M., Yan,S. and Patel,D.J. (2002) Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers*, **65**, 202–210.
77. Buterin,T., Meyer,C., Giese,B. and Naegeli,H. (2005) DNA quality control by conformational readout on the undamaged strand of the double helix. *Chem. Biol.*, **12**, 913–922.
78. Maillard,O., Camenisch,U., Clement,F.C., Blagoev,K.B. and Naegeli,H. (2007) DNA repair triggered by sensors of helical dynamics. *Trends Biochem. Sci.*, **32**, 494–499.
79. Chen,X., Velmurugu,Y., Zheng,G., Park,B., Shim,Y., Kim,Y., Liu,L., Van Houten,B., He,C., Ansari,A. *et al.* (2015) Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. *Nat. Commun.*, **6**, 5849.
80. Velmurugu,Y., Chen,X., Slogoff Sevilla,P., Min,J.H. and Ansari,A. (2016) Twist-open mechanism of DNA damage recognition by the Rad4/XPC nucleotide excision repair complex. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E2296–E2305.
81. Kong,M., Liu,L., Chen,X., Driscoll,K.I., Mao,P., Bohm,S., Kad,N.M., Watkins,S.C., Bernstein,K.A., Wyrick,J.J. *et al.* (2016) Single-nolecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process. *Mol. Cell*, **64**, 376–387.
82. Mu,H., Zhang,Y., Geacintov,N.E. and Broyde,S. (2018) Lesion sensing during initial binding by yeast XPC/Rad4: toward predicting resistance to nucleotide excision repair. *Chem. Res. Toxicol.*, **31**, 1260–1268.
83. Lin,G. and Li,L. (2010) Elucidation of spore-photoproduct formation by isotope labeling. *Angew. Chem. Int. Ed. Engl.*, **49**, 9926–9929.
84. Otwinowski,Z. and Minor,W. (1997) In: Carter,C.W. Jr (ed). *Methods Enzymol*. Academic Press, Vol. **276**, pp. 307–326.
85. Kabsch,W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 133–144.
86. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
87. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
88. The PyMOL Molecular Graphics System. (2018) Schrödinger, LLC. Version 2.1.1, https://pymol.org/.
89. Case,D.A., Betz,R.M., Cerutti,D.S., Cheatham,I.T.E., Darden,T.A., Duke,R.E., Giese,T.J., Gohlke,H., Goetz,A.W., Homeyer,N. *et al.* (2016) 16 edn. University of California, San Francisco, http://ambermd.org/.
90. 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.
91. Skjaerven,L., Yao,X.Q., Scarabelli,G. and Grant,B.J. (2014) Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinformatics*, **15**, 399.
92. Nairn,R.S., Mitchell,D.L., Adair,G.M., Thompson,L.H., Siciliano,M.J. and Humphrey,R.M. (1989) UV mutagenesis, cytotoxicity and split-dose recovery in a human-CHO cell hybrid having intermediate (6-4) photoproduct repair. *Mutat. Res.*, **217**, 193–201.
93. McCready,S. and Cox,B. (1993) Repair of 6-4 photoproducts in Saccharomyces cerevisiae. *Mutat. Res.*, **293**, 233–240.
94. McCready,S. (1994) Repair of 6-4 photoproducts and cyclobutane pyrimidine dimers in rad mutants of Saccharomyces cerevisiae. *Mutat. Res.*, **315**, 261–273.
95. Guzder,S.N., Sung,P., Prakash,L. and Prakash,S. (1998) Affinity of yeast nucleotide excision repair factor 2, consisting of the Rad4 and Rad23 proteins, for ultraviolet damaged DNA. *J. Biol. Chem.*, **273**, 31541–31546.
96. Jansen,L.E., Verhage,R.A. and Brouwer,J. (1998) Preferential binding of yeast Rad4.Rad23 complex to damaged DNA. *J. Biol. Chem.*, **273**, 33111–33114.
97. Chakraborty,S., Steinbach,P.J., Paul,D., Mu,H., Broyde,S., Min,J.H. and Ansari,A. (2018) Enhanced spontaneous DNA twisting/bending fluctuations unveiled by fluorescence lifetime distributions promote mismatch recognition by the Rad4 nucleotide excision repair complex. *Nucleic Acids Res.*, **46**, 1240–1255.
98. Yokoyama,H. and Mizutani,R. (2014) Structural biology of DNA (6-4) photoproducts formed by ultraviolet radiation and interactions with their binding proteins. *Int. J. Mol. Sci.*, **15**, 20321–20338.
99. Rycyna,R.E. and Alderfer,J.L. (1985) UV irradiation of nucleic acids: formation, purification and solution conformational analysis of the '6-4 lesion' of dTpdT. *Nucleic Acids Res.*, **13**, 5949–5963.
100. Taylor,J.S., Garrett,D.S. and Wang,M.J. (1988) Models for the solution state structure of the (6-4) photoproduct of thymidylyl-(3′—-5′)-thymidine derived via a distance- and angle-constrained conformation search procedure. *Biopolymers*, **27**, 1571–1593.
101. Yokoyama,H., Mizutani,R., Satow,Y., Komatsu,Y., Ohtsuka,E. and Nikaido,O. (2000) Crystal structure of the 64M-2 antibody Fab

fragment in complex with a DNA dT(6-4)T photoproduct formed by ultraviolet radiation. *J. Mol. Biol.*, **299**, 711–723.
102. Mu,H., Geacintov,N.E., Min,J.H., Zhang,Y. and Broyde,S. (2017) Nucleotide excision repair lesion-recognition protein Rad4 captures a pre-flipped partner base in a benzo[a]pyrene-derived DNA lesion: how structure impacts the binding pathway. *Chem. Res. Toxicol.*, **30**, 1344–1354.
103. Mu,H., Geacintov,N.E., Zhang,Y. and Broyde,S. (2015) Recognition of damaged DNA for nucleotide excision repair: a correlated motion mechanism with a mismatched cis-syn thymine dimer lesion. *Biochemistry*, **54**, 5263–5267.
104. Mu,H., Zhang,Y., Geacintov,N.E. and Broyde,S. (2018) Lesion sensing during initial binding by Yeast XPC/Rad4: toward predicting resistance to nucleotide excision repair. *Chem. Res. Toxicol.*, **31**, 1260–1268.
105. Song,K., Campbell,A.J., Bergonzo,C., de Los Santos,C., Grollman,A.P. and Simmerling,C. (2009) An improved reaction coordinate for nucleic acid base flipping studies. *J. Chem. Theory Comput.*, **5**, 3105–3113.
106. Rooklin,D., Wang,C., Katigbak,J., Arora,P.S. and Zhang,Y. (2015) AlphaSpace: fragment-centric topographical mapping to target protein-protein interaction interfaces. *J. Chem. Inf. Model.*, **55**, 1585–1599.
107. Maul,M.J., Barends,T.R., Glas,A.F., Cryle,M.J., Domratcheva,T., Schneider,S., Schlichting,I. and Carell,T. (2008) Crystal structure and mechanism of a DNA (6-4) photolyase. *Angew. Chem. Int. Ed. Engl.*, **47**, 10076–10080.
108. Yokoyama,H., Mizutani,R. and Satow,Y. (2013) Structure of a double-stranded DNA (6-4) photoproduct in complex with the 64M-5 antibody Fab. *Acta Crystallogr. D Biol. Crystallogr.*, **69**, 504–512.
109. Liebschner,D., Afonine,P.V., Moriarty,N.W., Poon,B.K., Sobolev,O.V., Terwilliger,T.C. and Adams,P.D. (2017) Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. D Struct. Biol.*, **73**, 148–157.
110. Cai,Y., Kropachev,K., Kolbanovskii,M., Kolbanovskii,A., Broyde,S., Patel,D. and Geacintov,N.E. (2010) In: Geacintov NE, and Broyde,S (eds). *The Chemical Biology of DNA Damage*. WILEY-VCH Verlag GmbH & Co., Weinheim, pp. 261–298.