

Lesion Sensing during Initial Binding by Yeast XPC/Rad4: Toward Predicting Resistance to Nucleotide Excision Repair

Hong Mu,[†] Yingkai Zhang,^{‡,§} Nicholas E. Geacintov,[‡] and Suse Broyde^{*,†,§}

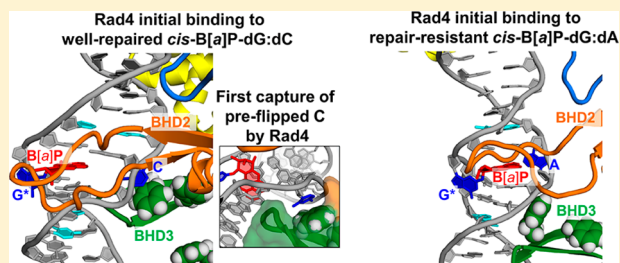
[†]Department of Biology and [‡]Department of Chemistry, New York University, New York, New York 10003, United States

[§]NYU-ECNU Center for Computational Chemistry at New York University Shanghai, Shanghai 200062, China

S Supporting Information

ABSTRACT: Nucleotide excision repair (NER) excises a variety of environmentally derived DNA lesions. However, NER efficiencies for structurally different DNA lesions can vary by orders of magnitude; yet the origin of this variance is poorly understood. Our goal is to develop computational strategies that predict and identify the most hazardous, repair-resistant lesions from the plethora of such adducts. In the present work, we are focusing on lesion recognition by the xeroderma pigmentosum C protein complex (XPC), the first and required step for the subsequent assembly of factors needed to produce successful

NER. We have performed molecular dynamics simulations to characterize the initial binding of Rad4, the yeast orthologue of human XPC, to a library of 10 different lesion-containing DNA duplexes derived from environmental carcinogens. These vary in lesion chemical structures and conformations in duplex DNA and exhibit a wide range of relative NER efficiencies from repair resistant to highly susceptible. We have determined a promising set of structural descriptors that characterize initial binding of Rad4 to lesions that are resistant to NER. Key initial binding requirements for successful recognition are absent in the repair-resistant cases: There is little or no duplex unwinding, very limited interaction between the β -hairpin domain 2 of Rad4 and the minor groove of the lesion-containing duplex, and no conformational capture of a base on the lesion partner strand. By contrast, these key binding features are present to different degrees in NER susceptible lesions and correlate to their relative NER efficiencies. Furthermore, we have gained molecular understanding of Rad4 initial binding as determined by the lesion structures in duplex DNA and how the initial binding relates to the repair efficiencies. The development of a computational strategy for identifying NER-resistant lesions is grounded in this molecular understanding of the lesion recognition mechanism.



INTRODUCTION

Nucleotide excision repair (NER) is a conserved defense mechanism against a broad range of pro-mutagenic DNA lesions derived from environmental pollutants, such as polycyclic aromatic chemicals.^{1–3} If these lesions are not repaired and progress to replication, they can cause mutations that initiate cancer. Hence, these repair-resistant lesions are the most detrimental ones to human health.⁴ While overall strategies of mammalian and prokaryotic NER have been elucidated, why the NER susceptibility of chemically different lesions varies greatly is not well understood. Hence, the problem of efficiently identifying the NER-resistant lesions from among the large number is an open challenge. The goal of the present study is to develop predictive computational strategies, bench-marked by experimental NER data, that will permit identification of repair-resistant lesions. Predictive toxicology is showing promise with the recently demonstrated capability for identifying toxicity of chemicals, such as candidate additives for foods, that rivals results obtained with animal testing.^{5,6}

There are two subpathways of NER, global genomic NER (GG-NER)^{1,7} and transcription-coupled NER (TC-NER),^{8–10} which differ only in the first lesion-recognition step. In TC-

NER, recognition is accomplished by a stalled RNA polymerase II. In GG-NER, the current focus, recognition is carried out by the XPC-RAD23B complex,² aided in cells by centrin²¹¹ and UV-DDB for UV photo lesions.^{12,13} UV-DDB is believed to hand off cyclobutane pyrimidine dimer (CPD) lesions to XPC.² This essential lesion recognition role of XPC has been well established by experimental studies, which have shown that XPC binding is required for the subsequent binding of the TFIIH complex.^{1,14–16} The XPD helicase in TFIIH verifies the lesion.¹⁷ Subsequently, other NER factors are recruited to ultimately produce excision of the 24–32-mer lesion-containing oligonucleotide, followed by repair synthesis to restore the DNA sequence.¹ Mutations in XPC cause a xeroderma pigmentosum disease that produces extreme UV sensitivity and skin cancers.¹⁸

The X-ray crystal structure of *Saccharomyces cerevisiae* Rad4-Rad23, a yeast orthologue of human XPC-RAD23B, bound to a CPD lesion-containing DNA duplex with mismatched partner thymines, has provided our first molecular level insight into lesion recognition by XPC¹⁹ (Figure 1). The trans-

Received: August 17, 2018

Published: October 4, 2018

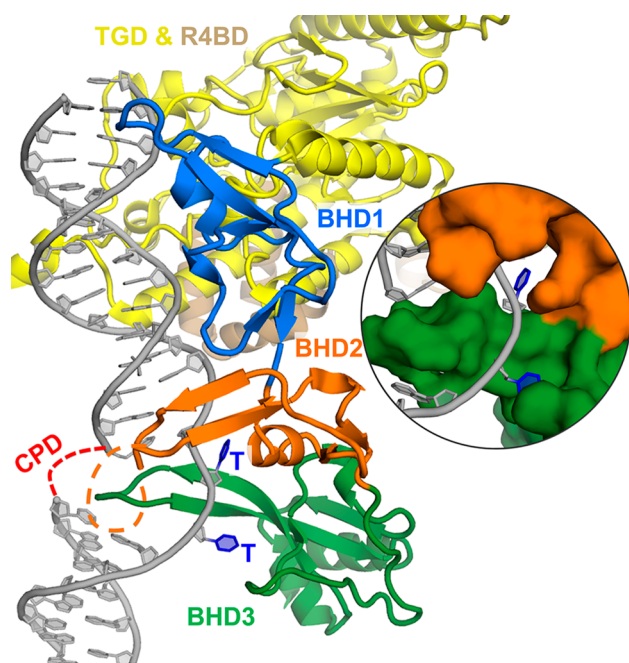


Figure 1. Crystal structure of the yeast orthologue of human XPC productively bound to CPD containing-DNA with mismatched thymines (PDB ID: 2QSG).¹⁹ The crystal structure is shown in cartoon representation. The TGD is yellow, the R4BD (Rad4/XPC binding domain in Rad23) is beige, BHD1 is marine, BHD2 is orange, BHD3 is dark green, and the DNA is light gray. The unresolved CPD (red) and BHD2 (orange) hairpin tip are indicated by dashed lines. The mismatched thymines (blue) that are flipped into their binding pockets are also shown in a zoomed-in view showing the surface of the binding pockets.

glutaminase domain (TGD) and β -hairpin domain 1 (BHD1) of Rad4 bind nonspecifically to the DNA sequence 3' to the lesion site, BHD2 binds at the lesion site from the minor groove side, and the β -hairpin domain 3 (BHD3) β -hairpin is inserted into the DNA helix at the lesion site from the major groove: The CPD is extruded, and its mismatched partner bases are flipped out and bound into their binding pockets in Rad4 (Figure 1). This structure, the productive open complex, suggests that β -hairpin insertion and flipping of the two partner strand bases are key elements of lesion recognition in NER. Since such productive binding requires denaturation at the lesion site, it is hypothesized that productive binding of XPC is facilitated by lesion-imposed DNA distortions and thermodynamic destabilizations, including especially ruptured Watson–Crick base pairing.^{20–24} However, little is known about how XPC binding varies for the diverse array of structurally different lesions.

For two different well-repaired lesion-containing duplexes, we have previously determined computationally the full binding pathways of Rad4/yeast XPC to the productive open complex.^{25,26} The two lesions were a small CPD in a duplex with mismatched partner thymines and a bulky 10R-(+)-*cis-anti*-benzo[*a*]pyrene-*N*²-dG (*cis*-B[*a*]P-dG) lesion in a duplex with normal partner base C. Differences in molecular details and energetics of the pathways were revealed, and differences are manifested upon the initial binding. Due to the dynamically extruded partner base C in the major groove of the *cis*-B[*a*]P-dG-containing duplex,^{27,28} the initial binding of Rad4 is quickly achieved via first capture of the partner C by the BHD2 and BHD3 domains. However, initial binding for the small CPD

with stacked-in partner thymines entailed probing of the minor groove by the BHD2 hairpin without partner base capture. Although different, the initial binding of Rad4 to both duplexes involves binding of BHD2 to the minor groove at the lesion site and distortions of the DNA duplex, especially unwinding. Here we hypothesize that initial binding of Rad4 varies for structurally different lesions with different NER susceptibilities and that poor XPC binding substrates would inhibit initial binding, leading to failure in lesion recognition and consequent repair resistance.

In the current study, we have explored, using extensive molecular dynamics (MD) simulations, Rad4/yeast XPC initial binding to a library of lesion-containing DNA duplexes that have different lesion topologies, stereochemistry, damaged base identity, and sequence context (Figure 2 and Figure S1). These lesions are derived from polycyclic aromatic chemicals that are mutagenic and carcinogenic environmental pollutants; their NER susceptibilities vary greatly, from resistant to well-repaired,^{22,28–31} as reviewed by Geacintov and Brodye.³² We quantified BHD2 binding into the minor groove using Alpha Space³³ and also quantified duplex untwisting/unwinding from its starting state prior to engagement of BHD2 at the lesion site. Our results showed that extent of BHD2 binding and unwinding of the duplex correlated well with NER efficiencies: Repair-resistant duplexes exhibit no significant unwinding with limited BHD2 interactions; only well-repaired duplexes show substantial unwinding and extensive BHD2 binding that is accompanied by conformational capture of a lesion partner base; and intermediate values for BHD2 binding and duplex unwinding are observed for moderately repaired lesions. Overall, these computational features of repair susceptibility and resistance are promising predictors of NER-resistant lesions based on our current lesion selection.

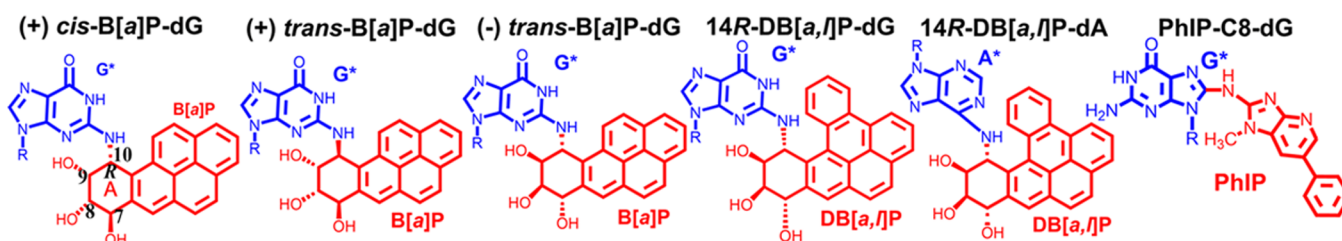
METHODS

The lesion containing 28-mer DNA duplexes were modeled based on known NMR/MD-derived structures (Figure 2 and Figure S1).^{22,27,28,34–39} Then an apo Rad4 model, based on the crystal structure of apo Rad4 (PDB ID: 2QSF), was docked on to the damaged DNA with its TGD and BHD1 domains positioned as in the crystal structure of the productive open complex (PDB ID: 2QSG). In these starting structures, BHD2 and BHD3 hairpins are close to the lesion site but not yet bound (Figure S1). MD simulations were performed for the starting protein–DNA complexes with Na⁺ counterions and explicit TIP3P water in a cubic periodic box using AMBER16.⁴⁰ The MD simulations for all cases were run to 1.5 μ s. Initial binding that was stable for ~ 0.5 μ s was achieved at ~ 1 μ s (Figures S2–S11), and the analyses were based on the ensemble between 1.0 and 1.5 μ s.

To evaluate duplex untwisting around the lesion site, we measured the twist angles for the lesion-containing 6-mer (Figure S1) using the cptraj module of AMBER16⁴⁰ and calculated the untwist angle at the initial binding state ($\text{Untwist} = \text{Twist}_{\text{start}} - \text{Twist}_{\text{initial binding state}}$) (Figure S1). $\text{Twist}_{\text{start}}$ is the ensemble average twist angle of the lesion-containing 6-mer during the first 1 ns of production MD, during which significant untwisting was not observed (Figures S2–S11); this ensemble represents the state of the lesion-containing sequence before the engagement of BHD2. $\text{Twist}_{\text{initial binding state}}$ is the twist angle of the lesion-containing 6-mer for the structures in the initial binding state ensemble (1–1.5 μ s). Positive values indicate further untwisting and negative values indicate further twisting.

To quantify BHD2's binding into the DNA minor groove, the best representative structure for the initial binding state of each lesion-containing duplex was analyzed. The best representative structure 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

A. Chemical structures



B. NMR/MD-derived structures

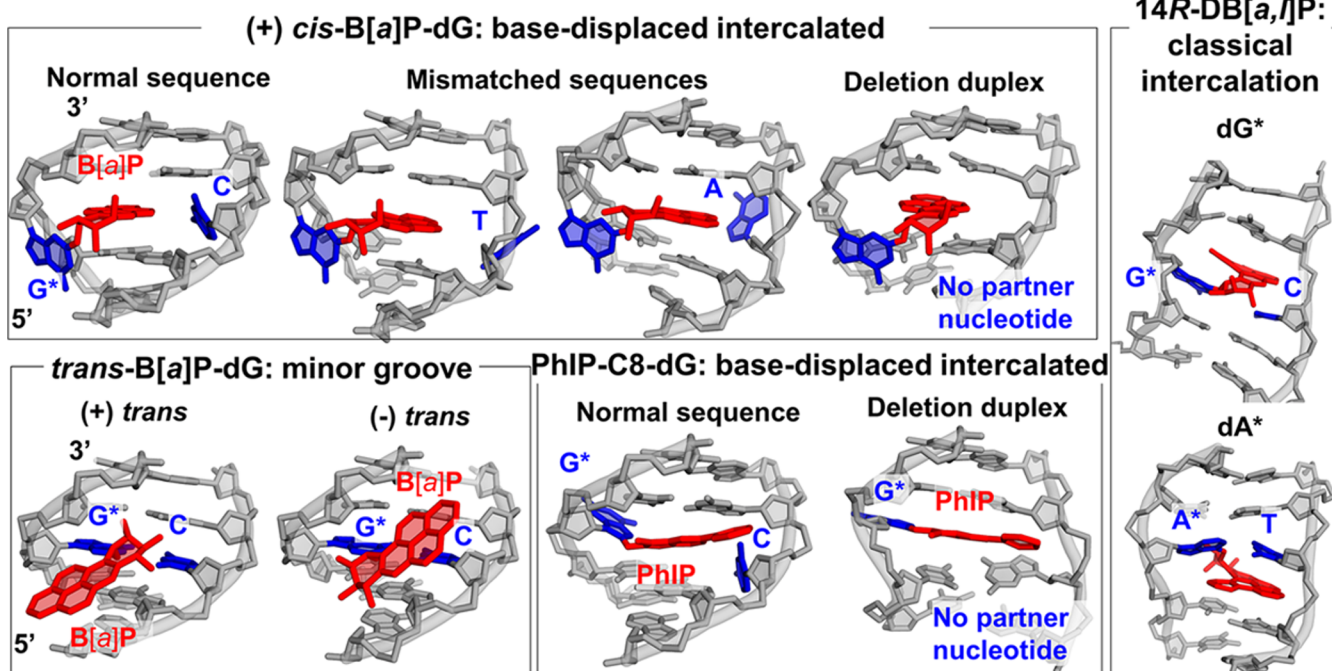


Figure 2. Lesion structures. (A) Chemical structures of the selected lesions: 10R-(+)-*cis-anti*-benzo[*a*]pyrene-*N*²-dG (*cis*-B[*a*]P-dG), 10S(+)-*trans-anti*-B[*a*]P-*N*²-dG [(+)-*trans*-B[*a*]P-dG], 10R(-)-*trans-anti*-B[*a*]P-*N*²-dG [(-)-*trans*-B[*a*]P-dG], 14R-(+)-*trans-anti*-dibenzo[*a,l*]pyrene-*N*²-dG (14R-DB[*a,l*]P-dG), 14R-(+)-*trans-anti*-dibenzo[*a,l*]pyrene-*N*⁶-dA (14R-DB[*a,l*]P-dA), and *N*-(deoxyguanosin-8-yl)-2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP-C8-dG). The benzylic ring is denoted by “A”. The * designates that the base is modified. (B) NMR/MD-derived structures of lesion-containing duplexes. The structures of the central 5-mers are shown in cartoon and sticks, with hydrogen atoms and backbone phosphate oxygens hidden. The lesion-containing base and its partner are blue, and the adduct is red. The view is into the minor groove. Full details for the NMR/MD-derived structures are reviewed in ref 32 and given in refs 22, 27, 28, and 34–39. The structural properties are fully summarized in Supporting Information. For the *cis*-B[*a*]P-dG cases, key differences entail the partner base identity and position or its absence. For the minor groove positioned *trans*-B[*a*]P-dG, the key difference is the 5′ vs 3′ orientation of the B[*a*]P ring system. For the PhIP-C8-dG, the mobile phenyl ring protruding in the minor groove is a key structural feature. For the bulky 14R-DB[*a,l*]P adducts, key differences are the intercalation from the minor groove with ruptured G*:C base pair for the dG* adduct but intercalation from the major groove with Watson–Crick pairing maintained for the dA* adduct.

of BHD2 to all other frames in each ensemble. The alpha space (V_α) volumes of the binding pockets in the DNA and their occupancies (O_α) by BHD2 were calculated using AlphaSpace v1.0.³³ The total occupied AS volume ($\sum V_\alpha \times O_\alpha$) was used to quantify the extent of BHD2 binding into the DNA minor groove (Figures S2–S11). The value reflects the curvature and surface area of the DNA minor groove region that is occupied by BHD2.

Full details on modeling, force field, MD protocols, and analyses are given in Supporting Information.

RESULTS

We carried out extensive MD simulations to explore initial binding states of Rad4 to a library of lesion-containing duplexes that were modeled based on their solution NMR/MD-derived structures (Figure 2B). These lesions differ greatly in chemical structures, conformations in duplex DNA, and

their experimental NER excision efficiencies that vary from resistant to efficiently repaired (Figure 3A). Initial binding states of Rad4 that were stable for $\sim 0.5 \mu\text{s}$ were obtained from structural ensembles in the 1.0–1.5 μs range of the production MDs. We devised computational characterizations that would reflect how well Rad4 binds initially to the lesion-containing DNA; these were based on experimental and computational studies of the XPC binding process, detailed in Discussion. These key descriptors are (1) duplex unwinding from the simulation starting conformation (Figure S1); (2) the BHD2 occupied AS volume³³ (Figures S2–S11), which reflects the curvature and surface area of the DNA minor groove that is bound by BHD2; and (3) lesion partner base flipping and capture by Rad4 (Figure 3).

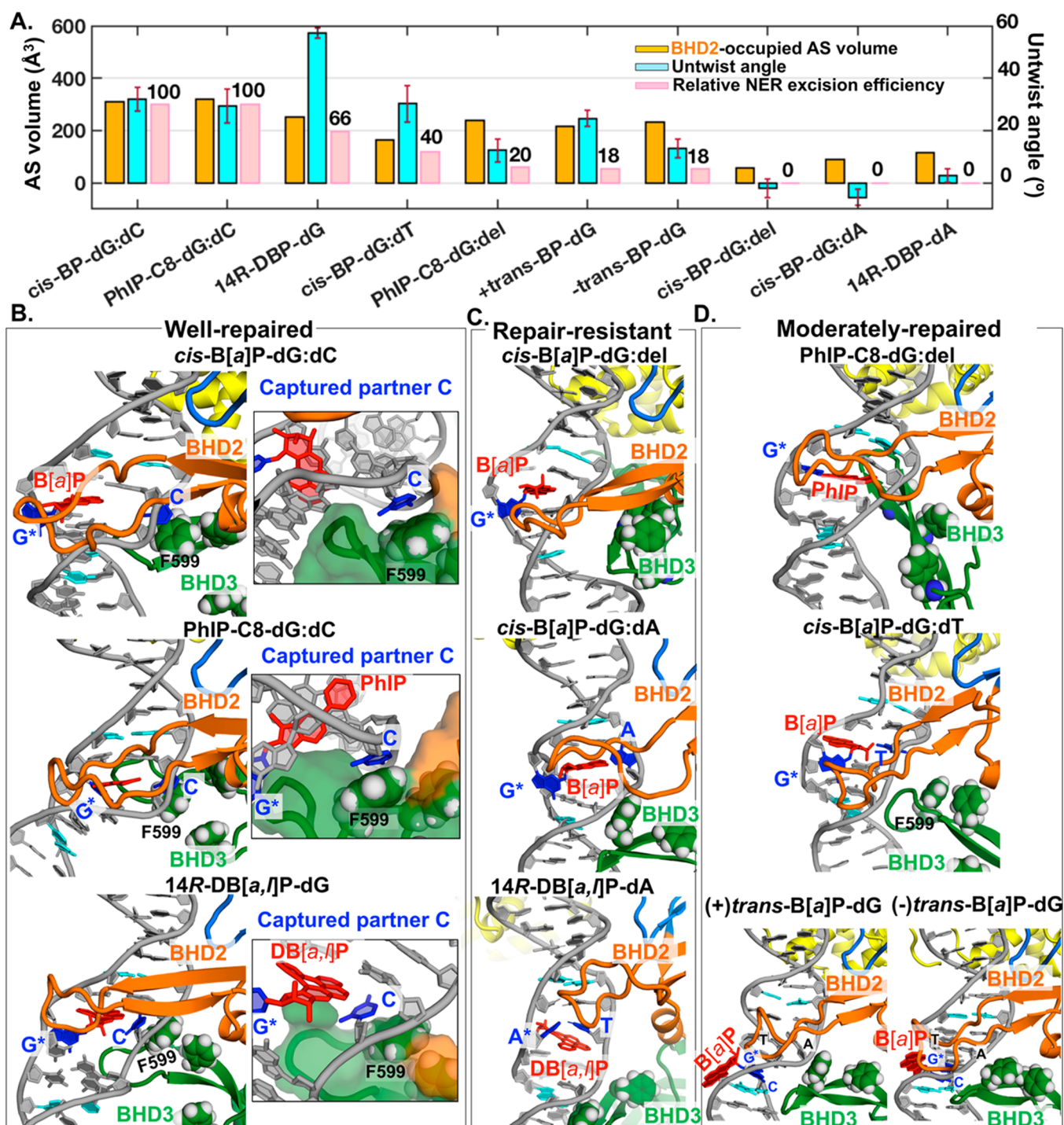


Figure 3. Structures and initial binding descriptors obtained from MD simulations and experimental NER excision efficiencies for the lesion-containing duplexes. (A) The AS volumes³³ occupied by BHD2 in the lesion-containing DNA are shown in orange bars; this volume reflects the curvature and surface area of the DNA minor groove bound by BHD2. The means and standard deviations for the block average values of untwist angles (detailed in Figure S1 and Supporting Information Methods) are shown in cyan bars and dark red lines. The relative NER excision efficiencies are in pink bars with the *cis*-B[a]P-dG:dC duplex assigned a relative value of 100. NER data are reviewed in ref 32 and are given in refs 22 and 28–31. (B–D) Best representative structures of the initial binding states from the MD trajectories. The structures are shown in cartoon with F556, F597, and F599 side chains in spheres. The base pairs used for calculation of untwist angles are in cyan. Insets depict zoomed-in views of captured partner strand base for the well-repaired cases, with BHD2 and BHD3 in surface representation. Movies S1–S10 show these initial binding state structures.

In the Well-Repaired Lesion-Containing Duplexes, the Lesion Partner Base Is Captured by the BHD3 Hairpin and BHD2 Binds Extensively, with Substantial Local Unwinding at the Lesion Site.

cis-B[a]P-dG:dC Duplex. We have further analyzed our previously obtained MD trajectory of Rad4, as it achieves the initial binding state with the base-displaced intercalated *cis*-B[a]P-dG-containing DNA duplex with normal partner C; this

is the first step in our study of the full binding pathway.²⁶ In this initial binding state, the preflipped partner C is captured in a pocket between BHD2 and BHD3 (Figure 3B), and BHD2 binds extensively into the minor groove around the lesion site. Here we analyzed the BHD2 occupied minor groove AS volume and obtained a high value of 309 Å³ (Figure S2). Correlated with this extensive insertion of BHD2 into the minor groove and partner base capture, our current analyses revealed that the DNA duplex is unwound around the lesion site with an untwist angle of 32° (Figure 3A).

PhIP-C8-dG:dC Duplex. At the initial binding state of Rad4 to the *N*-(deoxyguanosin-8-yl)-2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP-C8-dG)-containing duplex with normal partner C, the preflipped partner C in this base-displaced intercalated duplex is captured in a pocket between BHD2 and BHD3 (Figure 3B), and the BHD2 hairpin binds extensively in the minor groove at the lesion site, with a BHD2 occupied minor groove AS volume of 320 Å³ (Figure S3). Accompanying the insertion of the BHD2 hairpin into the minor groove, there is significant unwinding with an untwist angle of 29° (Figure 3A).

14R-DB[a,l]P-dG:dC Duplex. In this classically intercalated 14R-(+)-*trans-anti*-dibenzo[*a,l*]pyrene-*N*²-dG (14R-DB[*a,l*]P-dG)-containing duplex with normal partner C, the partner C is initially not Watson–Crick paired with the damaged guanine and extrudes during the binding process. A stable initial binding state is achieved when the partner C is captured by the pocket between BHD2 and BHD3 (Figure 3B). At the initial binding state, the BHD2 hairpin binds extensively in the minor groove, with a BHD2 occupied minor groove AS volume of 251 Å³ (Figure S4). The duplex unwinds with a very substantial untwist angle of 57° (Figure 3A), which results from the partner base flipping and its capture during the initial binding process (Figure S4).

Repair-Resistant Lesion-Containing Duplexes Do Not Exhibit Partner Base Capture, Achieve Very Limited Interactions with BHD2, and Do Not Exhibit Significant Duplex Unwinding.

***cis*-B[*a*]P-dG:Deletion Duplex.** We anticipated that initial binding of Rad4 would be inhibited by the missing partner nucleotide opposite the lesion and the strong van der Waals interactions between the B[*a*]P rings and adjacent base pairs, which would impede unwinding in this base-displaced intercalated duplex.²² Our results do show failed initial recognition of the lesioned duplex by Rad4. Rad4 stays stably at the lesion site for ~500 ns and then translocates one base pair step to the 3' side of the lesion-containing strand, where it remains stably for the last 500 ns (Figure S5). We investigated Rad4 initial binding at both of these locations. For the initial binding at the lesion site, the BHD2 hairpin does not insert into the minor groove (Figure 3C). It has limited interactions with the DNA backbone, with a very low BHD2 occupied minor groove AS volume of 58 Å³. Furthermore, there is no significant unwinding with an untwist angle of -2° (Figure 3A). For the initial binding at the translocated position 3' to the lesion, the results are similar, with a low BHD2 occupied minor groove AS volume of 124 Å³ and no unwinding of the DNA duplex (Figure S5).

***cis*-B[*a*]P-dG:dA Duplex.** The Rad4 initial binding to the *cis*-B[*a*]P-dG:dA duplex shows that extensive interactions of partner A with the major groove prevent the duplex from unwinding.²⁸ At the initial binding state, the BHD2 hairpin interacts via dynamic hydrogen bonds with the DNA

phosphate groups around the lesion site as well as by hydrophobic interactions with the lesion, but inserts only marginally (Figure 3C and Figure S6). The BHD2 occupied minor groove AS volume is a low value of 89 Å³ (Figure S6). There is no significant unwinding with an untwist angle of -6° (Figure 3A).

14R-DB[*a,l*]P-dA:dT Duplex. Upon Rad4 initial binding to the 14R-(+)-*trans-anti*-dibenzo[*a,l*]pyrene-*N*⁶-dA (14R-DB[*a,l*]P-dA)-containing duplex, there is no further unwinding, due to the strong van der Waals stacking interactions in the intercalation pocket of this classically intercalated structure with Watson–Crick pairing maintained.²³ At the initial binding state, the untwist angle is 3° (Figure 3A). BHD2 binds to a limited extent to the minor groove on the 3' side of the bulky DB[*a,l*]P rings, because the rings protrude to the minor groove and widen it (Figure 3C). The BHD2 occupied minor groove AS volume is 117 Å³ (Figure S7).

In the Moderately Repaired Lesion-Containing Duplexes, Partner Base Capture Does Not Occur, and BHD2 Binds Moderately, with Variable Local Unwinding at the Lesion Site.

PhIP-C8-dG:Deletion Duplex. Despite the absence of the partner nucleotide, this base-displaced intercalated duplex does manifest modest repair; this is in contrast to the *cis*-B[*a*]P-dG:deletion duplex which is repair-resistant and inhibits BHD2 binding and correlated unwinding. By comparison, the PhIP adduct is much less stable at the lesion site, with weaker stacking by the intercalated ring system, due to its topological difference: the mobile phenyl ring and smaller aromatic ring system. Upon initial binding by Rad4, this duplex can unwind modestly with an untwist angle of 12° and can open the minor groove to achieve a BHD2 occupied minor groove AS volume of 238 Å³ (Figure 3 and Figure S8). BHD2 in the minor groove engages in hydrophobic interactions with the PhIP phenyl ring. Two lesion-specific hydrogen bonds between the PhIP imidazole ring and Arg 601 at the tip of the BHD3 hairpin also stabilize the initial binding state (Figure S8).

***cis*-B[*a*]P-dG:dT Duplex.** Upon initial binding of Rad4, the mismatched partner T in this base-displaced intercalated structure becomes stacked into the duplex from its major groove position in the starting state;²⁸ it does not extrude into the major groove and does not become captured (Figure 3D). At the initial binding state, the BHD2 hairpin achieves moderate interactions with the minor groove at the lesion site, with a BHD2 occupied minor groove AS volume of 163 Å³ (Figure S9). In addition, the duplex unwinds with an untwist angle of 30° (Figure 3A).

(+)-*trans*-B[*a*]P-dG:dC Duplex. The minor groove position of the bulky aromatic lesion, 10S-(+)-*trans-anti*-B[*a*]P-*N*²-dG [(+)-*trans*-B[*a*]P-dG], obstructs insertion of the BHD2 hairpin into the minor groove at and 5' to the lesion site, leaving only the 3' side of the lesion site available for access by the BHD2 hairpin. This is consistent with experimental permanganate foot printing of human XPC binding to B[*a*]P-derived guanine lesions, which indicates that the binding sites of XPC in 10S-(+)-*trans*-B[*a*]P-dG and 10R-(-)-*trans-anti*-B[*a*]P-*N*²-dG [(-)-*trans*-B[*a*]P-dG]-containing duplexes are shifted to the 3' side of the lesion-containing strand.³⁰ We explored here the binding of Rad4 to the (+)-*trans*-B[*a*]P-dG:dC duplex starting from two positions: at the lesion site (Model 1) and at a T:A base pair two steps 3' to the lesion site where there is no obstruction (Model 2) (Figure S1). For Model 1, the incoming of BHD2 is blocked by the B[*a*]P rings, and the Rad4 quickly

translocates one step to the 3' side of the lesion-containing strand (Figure S10). In Model 2, this blockage is completely avoided, and stable initial binding of Rad4 is achieved: As Rad4 approaches, the BHD2 hairpin dynamically probes and ruptures the T:A base pair, which remains stacked-in but allows the DNA duplex to unwind (Figure 3D). BHD2 binds stably to the minor groove around this base pair with a BHD2 occupied minor groove AS volume of 216 Å³ (Figure S10). The duplex unwinds with an untwist angle of 25° (Figure 3A).

(-)-trans-B[a]P-dG:dC Duplex. Similar to the (+)-trans case, we explored the binding of Rad4 to the minor groove of the (-)-trans-B[a]P-dG:dC duplex starting from two positions: at the lesion site (Model 1) and at the T:A base pair two steps 3' to the lesion site (Model 2) (Figure S1). For Model 1, the insertion of BHD2 into the minor groove is blocked by the B[a]P rings (Figure S11). Stable binding of Rad4 is achieved for Model 2, where the obstruction is avoided. At the initial binding state, in contrast to the (+)-trans case, the T:A base pair remains intact (Figure 3D). The BHD2 hairpin binds moderately into the minor groove 3' to the lesion, with a BHD2 occupied minor groove AS volume of 232 Å³ (Figure S11). In concert, there is modest unwinding with an untwist angle of 13° (Figure 3A).

Portraits of stabilizing hydrogen-bonding and hydrophobic interactions between BHD2 and the DNA minor groove and backbone including the lesion itself, together with other details concerning Rad4 initial binding for each lesion, are given in Figures S2–S11. Movies S1–S10 show initial binding states.

DISCUSSION

Our goal in the present work was to begin the development of computational techniques for predicting the NER resistance of DNA lesions by exploring initial binding of Rad4/yeast XPC to lesion-containing duplexes, using extensive MD simulations. We hypothesize that poor initial binding inhibits lesion recognition by Rad4 and hence the subsequent cascade of events in NER. We devised three descriptors that reflect key structural features of Rad4 initial binding to lesion-containing DNA: the extent of BHD2 binding into the minor groove, the degree of duplex unwinding, and the occurrence of the partner strand base capture. These were based on insights into the Rad4 initial binding process provided by experimental single molecule⁴¹ and temperature-jump perturbation spectroscopy (T-jump) studies^{42,43} and computational pathway studies^{25,26} of Rad4 binding.

The existence of a fast initial DNA distortion step (~100–500 μs) was revealed in experimental studies of the Rad4 binding process using T-jump combined with fluorescence resonance energy transfer methods.⁴³ These studies suggested a “twist-open” mechanism for Rad4 binding: The initial binding causes fast DNA unwinding/untwisting, then the productive binding entails slow, full duplex opening (5–10 ms). It is proposed that the rates of these steps vary in different lesions due to their varying impacts on DNA distortion and destabilization. The existence of an initial binding step is further supported by single molecule studies.⁴¹

The importance of the Rad4 initial binding step is further supported by the kinetic gating mechanism: Lesion recognition is determined by a kinetic gate in which there is competition between Rad4 residence time at the lesion site and the time for Rad4 to achieve productive binding; and recognizable lesions allow Rad4 to reside at the lesion site long enough to form the

open complex, while Rad4 residence time is too short for the ones that escape recognition.⁴²

Furthermore, our computational pathway studies of Rad4^{25,26} showed that the initial binding varies for two different well-repaired lesions. It involves mainly BHD2 binding at the minor groove side of the lesion site, is enhanced by the capture of a partner strand base by the BHD2 and BHD3 domains, and is accompanied by DNA unwinding. The function of BHD2 in the initial lesion sensing is also supported by the T-jump and single molecule experiments, which showed that the BHD3 domain is not required for the initial binding stage.⁴³ In addition, deletion of a β-turn of BHD2 and BHD3 in human XPC greatly reduced XPC binding, but deletion of BHD3 alone had only modest impact on XPC binding to UV lesions in human cells.⁴⁴

Our quantification of extent of BHD2 binding into the minor groove and the degree of duplex unwinding upon initial binding shows remarkable correlation with experimental NER excision efficiencies for the 10 lesion-containing duplexes selected from our library, whose susceptibility varies from resistant to efficiently repaired (Figure 3A). With its diversity in lesion topology, stereochemistry, sequence context, and nature of adducted base (Figure 2), our selected lesions from the library reveal specific impacts on initial binding of Rad4. In particular, repair-resistant lesions show the least BHD2 binding and no unwinding, while more significant BHD2 binding and unwinding are observed for repair susceptible lesions (Figure 3). Furthermore, only the well-repaired duplexes manifest conformational capture of a partner base during initial binding (Figure 3B). Thus, these signature hallmarks of initial binding show promise as predictors of NER resistance, since failure to bind initially ensures that the following NER steps are impossible.

The role of the partner nucleotide is highlighted in our study of the repair-resistant *cis*-B[a]P-dG in a deletion duplex with a missing partner nucleotide. As expected, Rad4 fails to insert BHD2 and unwind the duplex at the initial binding state, consistent with experimental studies that show no specific binding of XPC for this case (Figure 3C). Nonetheless, in the PhIP-C8-dG deletion duplex, there is modest repair by NER despite the absence of the partner nucleotide. The unique topology of the PhIP lesion with its mobile phenyl ring and comparatively small ring system provides a less stable intercalation site structure, which can be somewhat unwound and accommodate BHD2 to an extent in the minor groove (Figure 3D). Furthermore, its heterocyclic chemical structure affords unique hydrogen bonds, and the phenyl ring provides hydrophobic interactions to Rad4 (Figure S8). We speculate that a neighboring partner strand base could be captured to achieve sufficient productive binding for modest NER.

Conformational capture of a partner base is also inhibited for the case of the *cis*-B[a]P-dG with mismatched partner A, which is repair resistant. Although displaced into the major groove, the mismatched A exhibits strong van der Waals interactions with the groove, which prevents its extrusion, duplex unwinding, and BHD2 binding to the minor groove, and explains its NER resistance (Figure 3C).

Another example of the importance of partner base conformational capture is revealed for the 14*R*-DB[*a,l*]P-adducted guanine and adenine pair. With intercalation of the DB[*a,l*]P ring system from the minor and major grooves, respectively, Watson–Crick pairing is ruptured only in the well-repaired guanine adduct, but maintained for the repair-

resistant adenine adduct (Figure 2). Strong stacking interactions between the bulky DB[*a,l*]P ring system and the adjacent base pairs in the 14R-DB[*a,l*]P-dA:dT duplex inhibit duplex unwinding and partner T flipping and limit BHD2 binding (Figure 3C). On the other hand, for the guanine adduct, the partner C to the modified guanine flips out upon BHD2 binding from its original stacked-in position and is captured by Rad4, accompanied by substantial unwinding (Figure 3B and Figure S4).

The modest repair of the minor groove positioned (+)- and (−)-*trans*-B[*a*]P-dG adducts is likely to be related to the obstruction of the minor groove for BHD2 binding at the lesion site. Our study reveals that BHD2 is able to achieve moderate binding and duplex unwinding two base pair steps 3′ to the lesion, where it avoids the obstruction of the minor groove by the lesion (Figure 3 and Figures S10–11).

We appreciate that the moderately repaired lesions, none of which display partner base conformational capture (Figure 3D), would eventually flip partner strand base/bases into Rad4 along the pathway toward productive binding. In order to improve our computational strategy for identifying repair-resistant lesions, further studies will be needed with additional selections of lesions and longer simulations to gain better understanding of Rad4 initial binding, including partner base flipping.

However, lesions may bind well initially but fail to achieve the productive open complex: BHD3 must insert into the duplex, which is required for recruitment of TFIIH needed for lesion verification and subsequent stages of NER. Hence, the current approach will identify only a subset of repair-resistant lesions, while full pathway studies with binding free energy barriers, as in our earlier work,^{25,26} would be needed to identify lesions that resist BHD3 insertion. Experimental XPC binding studies^{45,46} as well as T-jump^{42,43} and single molecule⁴¹ investigations together with ongoing NER characterizations of lesions provide anchors for the development of advanced computational strategies. In addition, the molecular and dynamic characterization of lesion verification remains obscure, and the role that verification may play in identifying repair-resistant lesions is a future challenge.

CONCLUSION

We have delineated promising structural predictors based on all-atom MD simulations of initial binding by Rad4/yeast XPC to a diverse array of DNA lesions with varying NER susceptibilities and structural features. Our descriptors characterize DNA lesions that are not sensed by Rad4/yeast XPC and hence are NER resistant and can persist to initiate cancer. Overall, our results show that repair-resistant lesions manifest little or no duplex unwinding upon Rad4 initial binding, very limited interaction between the BHD2 domain and the minor groove of the lesion-containing duplex, and no conformational capture of a base on the lesion partner strand. By contrast, the well-repaired lesions show extensive binding of BHD2, substantial unwinding, and conformational capture of a flipped partner base, while the moderately repaired lesions show intermediate binding of BHD2, varying levels of unwinding, and no conformational capture of a partner strand base. These differences are governed by the diverse structures of the lesions in duplex DNA. The characterizations provide molecular understanding of Rad4 initial binding and its relationship to repair efficiencies. Moreover, the identification of repair-resistant environmental carcinogens by computational

approaches is a timely endeavor, which will mitigate laborious experimental strategies that can include animal testing.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.chemrestox.8b00231.

MD simulation initial models, DNA base sequences and definition of twist and untwist angles (Figure S1), the initial binding of Rad4 to the *cis*-B[*a*]P-dG lesion site with preflipped partner C (Figure S2), the initial binding of Rad4 to the PhIP-C8-dG lesion site with preflipped partner C (Figure S3), the initial binding of Rad4 to the 14R-DB[*a,l*]P-dG lesion site with ruptured partner C (Figure S4), the initial binding of Rad4 to the *cis*-B[*a*]P-dG in the deletion duplex with its partner nucleotide missing (Figure S5), the initial binding of Rad4 to the *cis*-B[*a*]P-dG lesion site with major groove contacting mismatched partner A (Figure S6), the initial binding of Rad4 to the 14R-DB[*a,l*]P-dA lesion site with Watson–Crick paired partner T (Figure S7), the initial binding of Rad4 to the PhIP-C8-dG in the deletion duplex with its partner nucleotide missing (Figure S8), the initial binding of Rad4 to the *cis*-B[*a*]P-dG lesion site with major groove-contacting mismatched partner T (Figure S9), the initial binding of Rad4 to the (+)-*trans*-B[*a*]P-dG-containing duplex (Figure S10), the initial binding of Rad4 to the (−)-*trans*-B[*a*]P-dG-containing duplex (Figure S11), supplementary methods, movie descriptions (PDF)

Movie S1 (AVI)

Movie S2 (AVI)

Movie S3 (AVI)

Movie S4 (AVI)

Movie S5 (AVI)

Movie S6 (AVI)

Movie S7 (AVI)

Movie S8 (AVI)

Movie S9 (AVI)

Movie S10 (AVI)

AUTHOR INFORMATION

Corresponding Author

*Tel: +1 212 998 8231. E-mail: broyde@nyu.edu.

ORCID

Yingkai Zhang: 0000-0002-4984-3354

Suse Broyde: 0000-0002-3802-7511

Funding

This research was supported by National Institutes of Health (R01-ES025987 to S.B., R35-GM127040 to Y.Z., and R01-ES024050 to N.E.G.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation (NSF) grant MCB060037, and the high-performance computing resources of New York University (NYU-ITS). We thank Orlando Schärer (Distin-

guished Professor, Ulsan National Institute of Science and Technology) for reading the manuscript and providing very helpful suggestions.

■ ABBREVIATIONS

NER, nucleotide excision repair; XPC, xeroderma pigmentosum C protein-RAD23B complex; MD, molecular dynamics; BHD2, β -hairpin domain 2; GG-NER, global genomic NER; TC-NER, transcription-coupled NER; CPD, cyclobutane pyrimidine dimer; TGD, transglutaminase domain; BHD1, β -hairpin domain 1; BHD3, β -hairpin domain 3; *cis*-B[a]P-dG, 10R-(+)-*cis*-anti-benzo[a]pyrene-N²-dG; AS, alpha space; PhIP-C8-dG, *N*-(deoxyguanosin-8-yl)-2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine; 14R-DB[a,l]P-dG, 14R-(+)-*trans*-anti-dibenzo[a,l]pyrene-N²-dG; B[a]P, benzo[a]pyrene; 14R-DB[a,l]P-dA, 14R-(+)-*trans*-anti-dibenzo[a,l]pyrene-N⁶-dA; (+)-*trans*-B[a]P-dG, 10S-(+)-*trans*-anti-B[a]P-N²-dG; (-)-*trans*-B[a]P-dG, 10R-(-)-*trans*-anti-B[a]P-N²-dG; DB[a,l]P, dibenzo[a,l]pyrene; PhIP, 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine; T-jump, temperature-jump perturbation spectroscopy

■ REFERENCES

- Schärer, O. D. (2013) Nucleotide excision repair in eukaryotes. *Cold Spring Harbor Perspect. Biol.* 5, a012609.
- Mu, H., Geacintov, N. E., Broyde, S., Yeo, J. E., and Schärer, O. D. (2018) Molecular basis for damage recognition and verification by XPC-RAD23B and TFIIH in nucleotide excision repair. *DNA Repair*, DOI: 10.1016/j.dnarep.2018.08.005.
- Kisker, C., Kuper, J., and Van Houten, B. (2013) Prokaryotic nucleotide excision repair. *Cold Spring Harbor Perspect. Biol.* 5, a012591.
- Marteijn, J. A., Lans, H., Vermeulen, W., and Hoeijmakers, J. H. (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.* 15, 465–481.
- Hartung, T. (2018) Rebooting the generally recognized as safe (GRAS) approach for food additive safety in the US. *ALTEX* 35, 3–25.
- Luechtefeld, T., Marsh, D., Rowlands, C., and Hartung, T. (2018) Machine learning of toxicological big data enables read-across structure activity relationships (RASAR) outperforming animal test reproducibility. *Toxicol. Sci.* 165, 198–212.
- Sugasawa, K. (2016) Molecular mechanisms of DNA damage recognition for mammalian nucleotide excision repair. *DNA Repair* 44, 110–117.
- Hanawalt, P. C., and Spivak, G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.* 9, 958–970.
- Vermeulen, W., and Foustier, M. (2013) Mammalian transcription-coupled excision repair. *Cold Spring Harbor Perspect. Biol.* 5, a012625.
- Spivak, G. (2016) Transcription-coupled repair: an update. *Arch. Toxicol.* 90, 2583–2594.
- Nishi, R., Okuda, Y., Watanabe, E., Mori, T., Iwai, S., Masutani, C., Sugawara, 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Mathieu, N., Kaczmarek, N., and Naegeli, H. (2010) Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proc. Natl. Acad. Sci. U. S. A.* 107, 17545–17550.
- Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T. (2005) *DNA Repair and Mutagenesis*, 2nd ed., ASM Press, Washington, DC.
- Min, J. H., and Pavletich, N. P. (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* 449, 570–575.
- 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.
- 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.
- Reeves, D. A., Mu, H., Kropachev, K., Cai, Y., Ding, S., Kolbanovskiy, A., Kolbanovskiy, M., Chen, Y., Krzeminski, J., Amin, S., Patel, D. J., Broyde, S., and Geacintov, N. E. (2011) Resistance of bulky DNA lesions to nucleotide excision repair can result from extensive aromatic lesion-base stacking interactions. *Nucleic Acids Res.* 39, 8752–8764.
- Cai, Y., Geacintov, N. E., and Broyde, S. (2012) Nucleotide excision repair efficiencies of bulky carcinogen-DNA adducts are governed by a balance between stabilizing and destabilizing interactions. *Biochemistry* 51, 1486–1499.
- 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.
- 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.
- 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.
- Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Ibanez, V., Luna, E., Harvey, R., Geacintov, N. E., Broyde, S., and Patel, D. J. (1993) Solution conformation of the (+)-*cis*-anti-[BP]dG adduct in a DNA duplex: intercalation of the covalently attached benzo[a]pyrenyl ring into the helix and displacement of the modified deoxyguanosine. *Biochemistry* 32, 4145–4155.
- Mu, H., Kropachev, K., Chen, Y., Zhang, H., Cai, Y., Geacintov, N. E., and Broyde, S. (2013) Role of structural and energetic factors in regulating repair of a bulky DNA lesion with different opposite partner bases. *Biochemistry* 52, 5517–5521.
- Hess, M. T., Gunz, D., Luneva, N., Geacintov, N. E., and Naegeli, H. (1997) Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes. *Mol. Cell Biol.* 17, 7069–7076.
- Mocquet, V., Kropachev, K., Kolbanovskiy, M., Kolbanovskiy, A., Tapias, A., Cai, Y., Broyde, S., Geacintov, N. E., and Egly, J. M. (2007) The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. *EMBO J.* 26, 2923–2932.

- (31) Kropachev, K., Kolbanovskiy, M., Liu, Z., Cai, Y., Zhang, L., Schwaid, A. G., Kolbanovskiy, A., Ding, S., Amin, S., Broyde, S., and Geacintov, N. E. (2013) Adenine-DNA adducts derived from the highly tumorigenic Dibenzo[*a,l*]pyrene are resistant to nucleotide excision repair while guanine adducts are not. *Chem. Res. Toxicol.* 26, 783–793.
- (32) Geacintov, N. E., and Broyde, S. (2017) Repair-Resistant DNA Lesions. *Chem. Res. Toxicol.* 30, 1517–1548.
- (33) Rooklin, D., Wang, C., Katigbak, J., Arora, P. S., and Zhang, Y. K. (2015) Alpha space: fragment-centric topographical mapping to target protein-protein interaction interfaces. *J. Chem. Inf. Model.* 55, 1585–1599.
- (34) Brown, K., Hingerty, B. E., Guenther, E. A., Krishnan, V. V., Broyde, S., Turteltaub, K. W., and Cosman, M. (2001) Solution structure of the 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine C8-deoxyguanosine adduct in duplex DNA. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8507–8512.
- (35) Tang, Y., Liu, Z., Ding, S., Lin, C. H., Cai, Y., Rodriguez, F. A., Sayer, J. M., Jerina, D. M., Amin, S., Broyde, S., and Geacintov, N. E. (2012) Nuclear magnetic resonance solution structure of an N(2)-guanine DNA adduct derived from the potent tumorigen dibenzo[*a,l*]pyrene: intercalation from the minor groove with ruptured Watson-Crick base pairing. *Biochemistry* 51, 9751–9762.
- (36) Cosman, M., Fiala, R., Hingerty, B. E., Amin, S., Geacintov, N. E., Broyde, S., and Patel, D. J. (1994) Solution conformation of the (+)-*cis-anti*-[BP]dG adduct opposite a deletion site in a DNA duplex: intercalation of the covalently attached benzo[*a*]pyrene into the helix with base displacement of the modified deoxyguanosine into the minor groove. *Biochemistry* 33, 11518–11527.
- (37) Cai, Y., Ding, S., Geacintov, N. E., and Broyde, S. (2011) Intercalative conformations of the 14R (+)- and 14S (–)-*trans-anti*-DB[*a,l*]P-N(6)-dA adducts: molecular modeling and MD simulations. *Chem. Res. Toxicol.* 24, 522–531.
- (38) Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Singh, S. B., Ibanez, V., Margulis, L. A., Live, D., Geacintov, N. E., Broyde, S., et al. (1992) Solution conformation of the major adduct between the carcinogen (+)-*anti*-benzo[*a*]pyrene diol epoxide and DNA. *Proc. Natl. Acad. Sci. U. S. A.* 89, 1914–1918.
- (39) de los Santos, C., Cosman, M., Hingerty, B. E., Ibanez, V., Margulis, L. A., Geacintov, N. E., Broyde, S., and Patel, D. J. (1992) Influence of benzo[*a*]pyrene diol epoxide chirality on solution conformations of DNA covalent adducts: the (–)-*trans-anti*-[BP]G × C adduct structure and comparison with the (+)-*trans-anti*-[BP]G × C enantiomer. *Biochemistry* 31, 5245–5252.
- (40) Case, D. A., Betz, R. M., Cerutti, D. S., Cheatham, T.E., III, Darden, T. A., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., Lin, C., Luchko, T., Luo, R., Madej, B., Mermelstein, D., Merz, K. M., Monard, G., Nguyen, H., Nguyen, H. T., Omelyan, I., Onufriev, A., Roe, D. R., Roitberg, A., Sagui, C., Simmerling, C. L., Botello-Smith, W. M., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., Xiao, L., and Kollman, P. A. (2016) *AMBER 16*, University of California, San Francisco, San Francisco, CA.
- (41) 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., Min, J. H., and Van Houten, B. (2016) Single-molecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process. *Mol. Cell* 64, 376–387.
- (42) Chen, X., Velmurugu, Y., Zheng, G., Park, B., Shim, Y., Kim, Y., Liu, L., Van Houten, B., He, C., Ansari, A., and Min, J. H. (2015) Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. *Nat. Commun.* 6, 5849.
- (43) 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, E2296–2305.
- (44) Camenisch, U., Trautlein, D., Clement, F. C., Fei, J., Leitenstorfer, A., Ferrando-May, E., and Naegeli, H. (2009) Two-stage dynamic DNA quality check by xeroderma pigmentosum group C protein. *EMBO J.* 28, 2387–2399.
- (45) Lee, Y. C., Cai, Y., Mu, H., Broyde, S., Amin, S., Chen, X., Min, J. H., and Geacintov, N. E. (2014) The relationships between XPC binding to conformationally diverse DNA adducts and their excision by the human NER system: is there a correlation? *DNA Repair* 19, 55–63.
- (46) Feher, K. M. (2018) Characterization of xeroderma pigmentosum group C recognition of bulky DNA adducts and the correlation with nucleotide excision repair. Ph.D. Thesis, Department of Chemistry, New York University, New York.