HHS Public Access

Author manuscript

RSC Chem Biol. Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

RSC Chem Biol. 2021 April 1; 2(2): 523–536. doi:10.1039/d0cb00192a.

Light-induced modulation of DNA recognition by the Rad4/XPC damage sensor protein

Amirrasoul Tavakoli[1], **Debamita Paul**[1], **Hong Mu**[2], **Jagannath Kuchlyan**[1,‡], **Saroj Baral**[3], **Anjum Ansari**[3], **Suse Broyde**[2], **Jung-Hyun Min**[1,†]

[1]Department of Chemistry and Biochemistry, Baylor University, Waco, TX 76706, USA

[2]Department of Biology, New York University, New York, NY 10003, USA

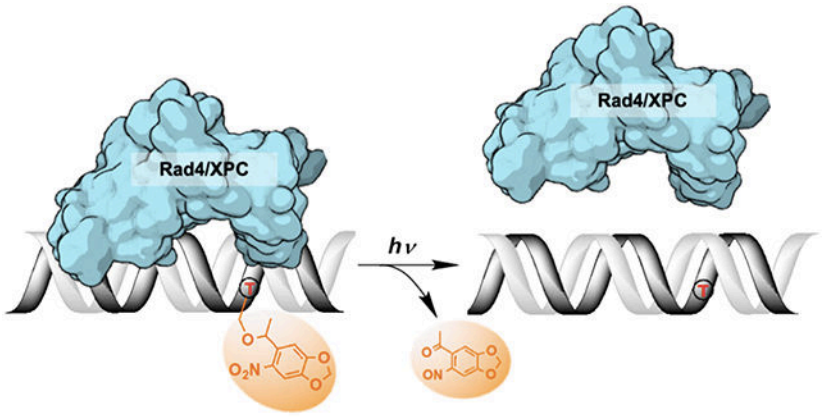
[3]Department of Physics, University of Illinois at Chicago, Chicago, IL 60607, USA

Abstract

Biomolecular structural changes upon binding/unbinding are key to their functions. However, characterization of such dynamical processes is difficult as it requires ways to rapidly and specifically trigger the assembly/disassembly as well as ways to monitor the resulting changes over time. Recently, various chemical strategies have been developed to use light to trigger changes in oligonucleotide structures, and thereby their activities. Here we report that photocleavable DNA can be used to modulate the DNA binding of the Rad4/XPC DNA repair complex using light. Rad4/XPC specifically recognizes diverse helix-destabilizing/distorting lesions including bulky organic adduct lesions and functions as a key initiator for the eukaryotic nucleotide excision repair (NER) pathway. We show that the 6-nitropiperonyloxymethyl (NPOM)-modified DNA is recognized by the Rad4 protein as a specific substrate and that the specific binding can be abolished by light-induced cleavage of the NPOM group from DNA in a dose-dependent manner. Fluorescence lifetime-based analyses of the DNA conformations suggest that free NPOM-DNA retains B-DNA-like conformations despite its bulky NPOM adduct, but Rad4-binding causes it to be heterogeneously distorted. Subsequent extensive conformational searches and molecular dynamics simulations demonstrate that NPOM in DNA can be housed in the major groove of the DNA, with stacking interactions among the nucleotide pairs remaining largely unperturbed and thus retaining overall B-DNA conformation. Our work suggests that photoactivable DNA may be used as a DNA lesion surrogate to study DNA repair mechanisms such as nucleotide excision repair.

Graphical Abstract

†To whom correspondence may be addressed, Tel: (254)710-2095, JungHyun_Min@baylor.edu.
‡Current address: Department of Chemistry, University of Oxford, Oxford, OX1 3TA, UK



Rad4/XPC DNA damage sensor protein specifically binds to photocleavable NPOM-DNA adduct, and this recognition is abolished upon photo-cleavage of NPOM.

Keywords

Photolabile DNA; photocage; phototrigger; photochemical conversion; optochemical control; light control; photo-cleavable; light-controlled protein-DNA interaction; photo-induced reaction; fluorescence lifetime; bulky adduct; DNA damage recognition; DNA damage repair; nucleotide excision repair; conformational dynamics; xeroderma pigmentosum; XPC; Rad4

INTRODUCTION

Biological processes entail dynamic yet coordinated assembly and disassembly of multiple molecules in solution. A key challenge in studying these processes in high structural and temporal resolution lies in the difficulties in controlled triggering of these events. Methods of triggers often involve either perturbing the equilibrium state using temperature-, salt- or pH-jumps or initiating the binding/unbinding through rapid mixing. Another method of triggering involves photo-induced chemical changes, as showcased by pioneering studies on dynamics of allosteric transitions in hemoglobin [1, 2]. Photoconvertible groups undergo structural changes upon irradiation by light, usually of a specific wavelength, in a reversible or irreversible manner [3, 4]. Such light-induced chemical/structural conversions have emerged as useful tools to control the properties and hence probe the functions of biomolecules harboring the photoconvertible groups, as light can be easily applied to biological systems *in vitro* or *in vivo* to trigger specific events [5-9]. Photoconvertible modifications in small molecules [10-12], oligonucleotides [13], peptides [14, 15], and proteins (mostly enzymes) [16-19] have also been used to control and monitor a wide variety of biological outcomes including gene expression, enzyme activity, oligomerization states, cellular localization and immune responses.

One of the photoconvertible groups, the 6-nitropiperonyloxymethyl (NPOM) has been introduced by the Deiters group as an improvement over previous *o*-nitrobenzyl derivatives for modifying biomolecules including oligonucleotides [20-22] (Figure 1). Irradiating an NPOM-modified nucleoside/ DNA/RNA with light (λ = 365 nm) cleaves the NPOM from the substrate without damaging the parent compound and restores the unmodified structures with concomitant release of 1-(6-nitroso-1,3-benzodioxol-5-yl)ethenone (hereafter

nitrosoacetophenone) (Figure 1) [23-26]. Compared with the previous O-nitrobenzyl derivatives such as 6-nitroveratryl (NV) and 6-nitroveratryloxycarbonyl (NVOC) groups, NPOM features a higher quantum yield ($\Phi = 0.094$ versus $\Phi = 0.0013$ for NVOC), is highly stable in an aqueous environment at various pHs [21, 23, 27] and penetrates cell membranes without altering growth rate or phenotype of the cells/organisms [28]. The NPOM modification has been applied to various *in vitro* and *in vivo* biological studies. For instance, photocleavage of NPOM-modified DNA/RNA has been used to control the activities of DNAzymes [29], antisense DNA/RNA [13, 30], restriction endonucleases [31], DNA-binding transcription factors [32], polymerase chain reaction (PCR) rates [33], as well as CRISPR-Cas gene editing [9, 18, 34, 35]. In these cases, NPOM-modifications prevent the DNA or RNA from hybridizing to the complementary strands, which could be reversed upon NPOM cleavage with light, converting the unhybridized 'inactive' molecules to hybridized 'active' molecules.

Here, we took advantage of the fact that DNA modifications such as NPOM are quite bulky and thus perhaps could be seen as a lesion on DNA by cellular DNA repair machineries, particularly those involved in the nucleotide excision repair (NER) pathway. NER repairs a wide spectrum of bulky adduct lesions in the DNA including sunlight-induced intra-strand crosslinks, bulky DNA adducts induced by various metabolites, reactive oxygen species, environmental pollutants and carcinogens (reviewed in [36-38]). Genetic impairment in NER causes high sun sensitivity xeroderma pigmentosum (XP) cancer predisposition syndrome in humans [37, 39]. In eukaryotic NER, the repair of these lesions scattered around the global genome is primarily initiated when the XPC–RAD23B complex (Rad4–Rad23 in yeast; hereafter referred to as XPC and Rad4) first specifically localizes to the lesion. The lesion binding by Rad4/XPC subsequently leads to the recruitment to the lesion of the transcription factor IIH complex (TFIIH) containing XPD and XPB helicases, which verifies the presence of a bulky lesion and recruits other NER factors. Eventually, a 24–32 nucleotide (nt) lesion-containing portion of the DNA strand is excised by the XPF-ERCC1 and XPG endonucleases and the gap in the DNA is restored by repair synthesis and nick sealing.

Previous studies from our group have shown that Rad4 recognizes DNA lesions in an indirect manner: crystal structures of Rad4 bound to UV-lesions showed that the Rad4 flips two damage-containing nucleotide pairs out of the duplex with the damaged nucleotides flipped away from the protein, such that only the undamaged nucleotides on the complementary strand make direct contacts with the protein (the so-called 'open' conformation) [40, 41]. This and other studies pose a puzzle as to the mechanism of this indirect readout by Rad4 and the nature of the structural intermediates along the recognition trajectory. These missing key steps led us to ponder whether photoconvertible adducts could serve as model DNA lesions and if their photoconvertible characteristics could be used to trigger the binding/unbinding events in a precisely controlled manner for structural and functional studies. So far, photoconvertible DNA/RNA has been mostly used for cellular and genetic studies but not as much for probing the biochemical and structural mechanisms, let alone for investigating NER [9, 42].

Here, using fluorescence lifetime-based Förster resonance energy transfer (FLT-FRET) measurements with cytosine analog FRET pair, tC^o and tC_{nitro}, that are uniquely sensitive to

changes in B-DNA conformation [43-46], together with UV-visible spectroscopy and competitive electrophoretic mobility shift assays (EMSA), we show that (1) the NPOM modification on DNA, though largely retaining B-DNA form, is specifically recognized by Rad4, (2) this specific binding is accompanied by heterogeneous structural distortions in DNA and (3) that this specific binding is abolished by photocleavage of the NPOM moiety from the DNA in a light-dose-dependent manner. Extensive conformational searches and molecular dynamics (MD) simulations of the DNA also corroborate with the fluorescence-based conformational analyses: the results demonstrate that the bulky NPOM moiety can be housed in the major groove of the DNA, with stacking interactions among the nucleotide pairs remaining largely unperturbed and thus retaining overall B-DNA conformation. Our findings provide the foundation for using NPOM and potentially other related photoconvertible DNA as novel probes to examine the damage recognition and repair mechanisms in the NER pathway. This groundwork opens doors to a variety of exciting future studies by coupling optical triggering with various techniques (e.g., fluorescence conformational dynamics spectroscopy, time-resolved x-ray crystallography, cellular repair kinetic studies) and contributes to expanding the current applicability of photochemistry to the DNA repair field.

MATERIALS AND METHODS

Preparation of Rad4–Rad23 complexes.

The Rad4-Rad23 complex (or simply referred to as Rad4) was prepared as previously described [40, 41, 47]. The Rad4 construct spans residues 101–632 and contains all four domains involved in DNA binding. This Rad4-Rad23 construct has previously been shown to exhibit the same DNA-binding characteristics as the full-length complex [41]. While Rad23 does not participate in DNA binding directly, it is required for stabilizing Rad4.

Hi5 insect cells co-expressing Rad4 and Rad23 proteins were harvested 2 days after infection. After lysis, the protein complex was purified by affinity chromatography (Ni-NTA Agarose, MCLAB), anion-exchange (Source Q, GE healthcare) and cation exchange (Source S, GE healthcare) chromatography followed by gel filtration (Superdex200, GE healthcare). The chromatogram and SDS-PAGE analyses of the gel filtration step show that peak fractions contain a homogeneous 1:1 complex of Rad4 and Rad23 proteins. These peak fractions were pooled and further concentrated by ultrafiltration (Amicon Ultra-15, Millipore) to ~13–14 mg/ml (135–150 μM) in 5 mM bis-tris propane–HCl (BTP-HCl), 800 mM sodium chloride (NaCl) and 5 mM dithiothreitol (DTT), pH 6.8. The complex was prepared without thrombin digestion, thus retaining the UBL domain of Rad23 and a histidine-tag on Rad4.

Preparation of double-stranded DNA substrates.

Unmodified DNA oligonucleotides were purchased from Biosynthesis or Integrated DNA Technologies (IDT). DNA oligomers synthesized with tC^o and tC_{nitro} were from Biosynthesis. All oligonucleotides were purchased as HPLC-purified. Oligonucleotides appeared as a single band on denaturing polyacrylamide gels, indicating high length-purity (>90%) of the oligonucleotides. The concentrations of each single-stranded DNA were

determined by UV absorbance at 260 nm (NanoDrop One^C, ThermoFisher) using the A_{260} extinction coefficients calculated by the nearest neighbor method (Biosynthesis). To prepare DNA duplexes, two complementary oligonucleotides were mixed at 1:1 molar ratio at 100 μM in TE buffer (10 mM Tris-HCl pH 7.50, 1 mM EDTA) in a microcentrifuge tube and annealed by slow-cooling: the tube was immersed in a 1.2 L hot water bath (~100 °C) placed on a hot plate; after 5 minutes, the hot plate was turned off and the samples were cooled to room temperature over 5 to 6 hours.

Photo-irradiation experiments and cleavage studies with NPOM DNA and NPOM-dT (NPOM-modified deoxythymidine).

NPOM-dT was purchased from Berry & Associates (Cat No. PY 7795). 50 μl of 10-50 μM of DNA samples or NPOM-dT were irradiated in a chamber encasing four UV-A lamps, at 6.4 mW/cm^2 (measured by General Tools Digital UVA/UVB Meter, 280-400 nm; #UV513AB). At specified time intervals, 1.5 μl of the sample was taken out and its absorbance spectrum was measured with NanoDrop One UV-Vis Spectrophotometer. The experiments were done with lights off in the lab. Exposing the NPOM-dT 10-400 μM in TE buffer with 0.01-0.4% DMSO or NPOM-DNA 10-50 μM in TE buffer under ambient light in the lab for 24 hours did not change the absorbance spectra of the samples, indicating little photocleavage by ambient light. For irradiation at 405 nm (blue light), we used a 405 nm LED lamp strip (6 W total output with delivered power estimated as ~4-8 mW/cm^2). The progress of the photocleavage reaction was monitored using the UV-visible spectrophotometer as described above. The time courses of the photocleavage reactions were analyzed by single exponential fitting using OriginPro 9.7.0.188 software (OriginLab).

Melting temperature measurements of DNA duplexes.

The overall thermal stabilities of all DNA duplexes were measured as follows. The absorbance at 260 nm of each DNA duplex (1.5 μM) was measured in a sample cuvette of path length 1 cm, using Cary 300 Bio UV-Visible spectrophotometer equipped with a Varian temperature controller. The absorbance measurements were done from 10 to 85 °C at every 1.0 °C interval. Derivative method [48] of Carry300 (Thermal software) was used to calculate the melting temperature (T_m) at which 50% of the DNA strands have separated. The derivatives were obtained numerically from the absorbance data using a Savitzky Golay technique where the difference between adjacent points was first computed followed by a smoothing procedure where 5 points surrounding an individual point were averaged to produce a new, smoothed point [49].

Competition electrophoretic mobility shift assays (EMSA).

To determine the relative affinities of Rad4 binding to different DNA substrates, competition EMSA (or gel-shift assays) were employed essentially as previously described [40, 41, 43, 47, 50, 51]. The benefit of using this competition assay over the conventional single-substrate EMSA is that one can directly observe any preferential binding over the nonspecific binding, including factors such as potential DNA end-binding while avoiding multiple proteins aggregating on a single DNA, as is the case when protein is in excess of total DNA [51]. Various concentrations of the Rad4-Rad23 complexes were mixed with 5 nM ^{32}P-labelled DNA of interest (mismatched/damaged or matched/undamaged) in the presence

of 1000 nM cold (unlabeled), matched DNA, CH7_NX in an EMSA buffer (5 mM BTP-HCl, 75 mM NaCl, 5 mM DTT, 5% glycerol, 0.74 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 500 μg/ml bovine serum albumin, pH 6.8). Mixed samples were incubated at room temperature for 20 min and separated on 4.8% non-denaturing polyacrylamide gels in 1x TBE buffer (89 mM Tris-HCl, 89 mM boric acid, 2 mM EDTA, pH 8.0), run at constant 150 V for 15 min at 4 °C. The gels were quantitated by autoradiography using Typhoon FLA9000 and Imagelab 6.0.1 software (Bio-Rad). The averages of the Rad4-bound DNA fractions quantified from three independent EMSA gels were used for subsequent calculations of the apparent dissociation constants ($K_{d,app}$).

To obtain apparent dissociation constants ($K_{d,app}$) for different DNA substrates, we first used the matched CH7_NX DNA as both the 'hot' probe and the cold competitor DNA, and obtained the $K_{d,app}$ for CH7_NX (K_{ns}) by fitting the fraction of labelled DNA bound (f) to the equation

$$f^2 \cdot [D_{ns}]_t - f \cdot ([D_{ns}]_t + [P]_t + K_{ns}) + [P]_t = 0$$

where $[P]_t$ is the total protein concentration and $[D_{ns}]_t$ is the total CH7_NX concentration (1005 nM). The $K_{d,app}$'s of other DNA substrates (K_s) were subsequently obtained by using the DNA of interest as the ^{32}P-labeled DNA probe and fitting the fraction of labelled DNA bound (f) to the equation:

$$f^2 \cdot \{K_s^2 - ([D_{ns}]_t - [P]_t + K_{ns}) \cdot K_s - [P]_t \cdot K_{ns}\} + f \cdot \{([D_{ns}]_t - [P]_t + K_{ns}) \cdot K_s + 2 \cdot [P]_t \cdot K_{ns}\} - [P]_t \cdot K_{ns} = 0$$

where $[P]_t$ is the total protein concentration, $[D_{ns}]_t$ is the concentration of the undamaged competitor CH7_NX (1000 nM), and K_{ns} is the $K_{d,app}$ for CH7_NX binding, as obtained above. The equation for K_s was obtained using the approximation that the concentration of Rad4-bound labelled DNA is negligible compared to the total concentrations of Rad4 and of the matched/undamaged DNA competitor. Curve fittings for K_{ns} and K_s were both done by the nonlinear regression method using Origin software (OriginLab). The errors reported for $K_{d,app}$ indicate the errors of the nonlinear regression fit [40, 41].

Time-resolved fluorescence spectroscopy for fluorescence lifetime (FLT) measurements.

DNA duplexes labeled with both tC^o and tC_{nitro} (DNA_DA) or tC^o alone (DNA_D) were prepared as described above. The DNA and Rad4-Rad23-DNA 1:1 complex were prepared at 5 μM in phosphate-buffered saline (PBS: 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl pH 7.4) with 1 mM DTT. Under this condition, native gel electrophoresis and dynamic light scattering experiments showed that the Rad4-Rad23–DNA samples form uniformly sized 1:1 protein:DNA complexes [51]. Sample volume for each fluorescence lifetime (FLT) measurement was 12 μl. Fluorescence decay curves for the FRET donor tC^o (in the absence and presence of the FRET acceptor tC_{nitro}, which in itself is nonfluorescent) were measured with DeltaFlex single-photon counting instrument (HORIBA) equipped with a Ti-sapphire laser as the excitation light source (Mai Tai HP,

Spectra-Physics). The beam for tC^o excitation was produced by frequency doubling of the fundamental beam (730 nm) and pulse-picking at 4 MHz, which was then passed through a monochromator set at 365 nm (band pass 10 nm). The fluorescence signal emitted at 470 nm (band pass 10 nm) was collected by a Picosecond Photon Detection module (PPD-850, Horiba) using time-correlated single-photon counting (TCSPC) electronics. Fluorescence decay curves were recorded on a 100 ns timescale, resolved into 4,096 channels, to a total of 10,000 counts in the peak channel. All details are in Supplementary Methods.

Analysis of the fluorescence decay traces using discrete exponential (DE) fits.

The discrete exponential (DE) analysis was carried out using EzTime software (version 3.2.9.9, Horiba) that uses a standard iterative reconvolution method, assuming a multiexponential decay function, $I(t) = \sum_{i=1}^{n} \alpha_i \exp(-t/\tau_i)$, where α_i is the amplitude and τ_i is the fluorescence lifetime of the i-th decay component (Table S1). The maximum number of exponentials allowed by this software is five. For all measured decay traces, no more than four exponentials were needed to reasonably fit the data. The number of exponentials required for each trace was determined by the quality of the fit, evaluated based on the reduced chi-square χ^2 and the randomness of residuals (Figure S6). Each exponential component for the donor-acceptor labeled samples (DNA_DA) was characterized in terms of a lifetime denoted as $\tau_{DA,i}$, and a corresponding normalized amplitude or relative population $A_i = \frac{\alpha_i}{\sum_i \alpha_i}$. The FRET efficiency for the population in that component was computed from $E_i = 1 - \frac{\tau_{DA,i}}{\tau_D}$, where τ_D indicates the intrinsic lifetime of the donor probe. The average FRET efficiency for each sample was computed as $< E > = \sum_i A_i E_i = 1 - \frac{< \tau_{DA} >}{\tau_D}$, where $<\tau_{DA}>=\sum_i A_i \tau_{DA,i}$. For cases where the intrinsic lifetime of the donor-only samples could not be described by a single exponential, τ_D was taken as the intrinsic lifetime of the donor probe obtained from unmodified DNA (AT10_D).

Analysis of the decay traces using maximum entropy method (MEM) and Gaussian fitting.

Though the DE fitting has been traditionally used for fluorescence lifetime decay analyses, MEM has distinct advantages [43, 52]. In our previous studies, we have also shown that the results from the DE and MEM analyses corroborate with each another [43]. The MEM analyses were carried out using MemExp software [53, 54], as done previously [43]. The reproducibility of the distributions obtained from the MEM analyses from three independent lifetime measurements on each sample are illustrated in (Figure S8). The data presented in (Figure 4) are for one representative from this set. To further characterize the lifetime distributions from the MEM analyses, we fitted the measured distributions to a sum of Gaussians (Figure S7). Each Gaussian component was used to calculate the average FRET representing that component, and the area under the Gaussian curve was taken as a measure of the fractional population of that component. The results are summarized in Table S2. Errors are indicated with standard deviations (s.d.) from three independent sets of measurements.

Conformational searches and molecular dynamics simulations of NPOM-dT-containing DNA duplex structures.

In order to explore the structures of the NPOM-dT-containing DNA duplex, we first modeled NPOM-dT at the center of a 13-mer B-DNA duplex with the same sequence as in the AT2 NPOM-DNA (Table 1) employed in the experimental study. We carried out extensive conformational searches beginning with an NPOM-dT nucleoside to generate initial models for MD simulations of NPOM-dT in a B-DNA duplex, utilizing a sequence of protocols involving molecular modeling (Discovery Studio 2.5, Accelrys Software Inc.) and quantum mechanical geometry optimization (Gaussian 09[55]) to define sterically feasible NPOM-dT rotamer combinations for initiating the MD simulations (Figures S13-S14). These protocols and obtained structures are summarized in Scheme 1. We used the AMBER18 suite of programs [56] for MD simulations and analyses. Full details of NPOM-dT force field parameterization, MD simulation methods and analyses are given in the SI Methods. Newly developed force field parameters for the NPOM-dT are given in Table S3.

RESULTS

Dose-dependent photocleavage of NPOM from DNA as monitored by UV-visible absorption spectroscopy.

Photoconversion reactions often induce changes in the absorption spectra of the chemical groups of interest, which in turn can be used to track the reaction progress. To monitor the NPOM's photocleavage reaction, we obtained the UV-visible absorption spectra of the NPOM-modified DNA duplex (NPOM-DNA or AT2; see Table 1 for the DNA sequences used in the study) and dT nucleoside (NPOM-dT) after they were irradiated for varying time periods with photocleavage-inducing light (λ=365 nm). The overall absorbance in the 300-500 nm range increased with increased irradiation time, with the absorption maximum (λ_{max}) shifting from 365 nm to 395 nm (red-shift) for both samples, saturating by ~2-3 min time points (Figures 2 and S1). The time courses by which the absorbance at 395 nm reached saturation were similar for NPOM-DNA and NPOM-dT, accompanied by a common ~3-fold increase upon saturation (Figure 2A, B). Unmodified DNA duplex did not show absorption in this wavelength range with or without irradiation (Figure 2B). The strong absorbance at 395 nm after irradiation comes from the photocleavage reaction product, nitrosoacetophenone (Figure 1) [57]. When the nitrosoacetophenone was removed from DNA using a size-exclusion purification (G25, MWCO ~5 kDa), the absorption spectrum of the sample largely returned to that of the unmodified DNA (Figure S1D). Recently, works using visible light for release of *o*-nitrobenzyl or nitropiperonyl photocaging groups in DNA or RNA have been reported [58-60]. At the suggestion of a reviewer, we thus also tried 405 nm light as the light source and observed that it also induced efficient photocleavage of nitrosoacetophenone from the NPOM-containing substrates (Figure S2). Altogether, the results confirm the photo-induced cleavage of NPOM-DNA and indicate that the photocleavage reaction can be modulated by light doses [9]. The photocleavage reaction may, in principle, be accelerated by using light of higher intensity and shorter time duration, as indicated previously [9, 61]. Although there have been multiple studies using NPOM as photoactivatable group that elicit various biological outcomes, this is the first time the

photocleavage reaction progress was characterized *in situ* (through monitoring of the absorption spectra).

NPOM lowers the thermal stability of DNA duplex which can be reversed by photocleavage.

Several studies show that Rad4/XPC-binding and NER repair propensity for various lesions are positively correlated with the thermal destabilization induced by the lesion, which can be measured by DNA melting temperatures [40, 51, 62]. To see if thermal stability of DNA was impacted by the NPOM modification, we measured the melting temperatures (T_m) of the NPOM-DNA before and after photocleavage and compared them with that of the unmodified DNA. The T_m of NPOM-DNA (AT2, 45.2 °C) was ~7 °C lower than that of the unmodified DNA (AT1, 52.0 °C) while the T_m of NPOM-DNA after photocleavage (2 min irradiation) was the same as that of the unmodified DNA (52.0 °C) (Table 1, Figure S3A). These results showed that covalent NPOM adduct destabilized the DNA duplex but its photoremoval restored the DNA stability. The reaction products such as nitrosoacetophenone, though present in the reaction mix, did not affect the DNA thermal stability.

Competitive electrophoretic mobility shift assays (EMSA) show that NPOM-DNA is specifically recognized by Rad4, which is abolished upon NPOM photocleavage.

After observing NPOM is a helix-destabilizing DNA adduct, we set out to examine if the adduct can indeed be recognized as a DNA lesion by Rad4/XPC, by using a competitive electrophoretic mobility shift assay (EMSA) as extensively used before (Figure 3) [40, 41, 43, 47, 50, 51]. In this assay, the binding of the protein to 5 nM ^{32}P-labeled substrate DNA is monitored in the presence of 1000 nM unlabeled, undamaged “competitor” DNA (CH7_NX). The NPOM-DNA (AT2) showed ~15-fold lower apparent dissociation constant ($K_{d,app}$ ~48 nM) than the corresponding unmodified DNA (AT1) ($K_{d,app}$ ~701 nM). This specificity of NPOM-DNA (AT2) is even slightly higher than another specific model DNA substrate containing CCC/CCC mismatches (CH10_NX; $K_{d,app}$ ~79 nM), and is comparable to that of a bona fide NER lesion, 6-4 thymidine-thymidine photoproduct (6-4PP) ($K_{d,app}$ ~35 nM) [40]. On the other hand, the NPOM-DNA after photocleavage (AT2 + $h\nu$) showed $K_{d,app}$ (~744 nM) comparable to that of the unmodified DNA (AT1). These results show that the NPOM adduct in DNA is specifically recognized by Rad4/XPC as a lesion and that its photoremoval abolishes the specific binding. Additionally, when the protein and NPOM-DNA mixture was co-irradiated, the level of protein-bound NPOM-DNA progressively decreased with increasing irradiation time, consistent with the loss of specificity as NPOM is photo-cleaved from the DNA. The result also confirms that the photoirradiation did not induce erratic protein-DNA crosslinks (Figure S4). These results therefore reveal the potential of NPOM adduct to be used as a new model DNA lesion system whose specificity can be controlled by light, for studying Rad4/XPC and NER.

DNA conformation landscape mapped by fluorescence lifetime analyses of the tC^o-tC$_{nitro}$-labeled DNA shows NPOM-DNA becomes heterogeneously distorted upon binding to Rad4.

Previously, we showed that fluorescence lifetime (FLT) analyses combined with a set of FRET probes, tC^o and tC$_{nitro}$, in DNA can be used to map the conformations of DNA in solution [43]. The tC^o and tC$_{nitro}$ are a FRET pair that serve as donor and acceptor, respectively [44, 45]. As cytosine analogs, these probes retain normal Watson–Crick pairing with guanines with minimal perturbation of DNA structure and stability [44, 51]. Furthermore, the rigid exocyclic ring and its base stacking interactions hold these nucleotide analogs in relatively fixed orientations within the DNA helical structure, making their FRET sensitive to subtle distortions in DNA helicity that alter the probes' separation and/or relative orientation [63-65]. For example, Rad4-induced untwisting and 'opening' of 3-bp mismatched DNA could be monitored by the FRET efficiency between tC^o and tC$_{nitro}$ placed on either side of the mismatch [43, 51]. The FRET efficiency (E) relates directly to the lifetimes of the excited donor fluorophore, as $E = 1 - \frac{\tau_{DA}}{\tau_D}$, where τ_{DA} and τ_D are the donor lifetimes in the presence and absence of the acceptor, respectively. The lifetime approach offers distinct advantages over other techniques such as single-molecule FRET and is a more robust way to obtain FRET efficiency than the intensity-based steady-state measurements [43, 66]. Here, we adopted our previous approach and incorporated the tC^o-tC$_{nitro}$ FRET probes in the context of the NPOM-DNA construct (AT2) in the same positions relative to the lesion site as before (Table 1)[43, 51]. As expected, tC^o-tC$_{nitro}$ probes did not significantly alter the overall DNA duplex stability (as measured by T_m) of these constructs (Figure S3B) [43, 51]. Next, the fluorescence decays of each sample were obtained (Figure S5) and analyzed using two different methods, discrete exponential (DE) and maximum entropy method (MEM), as before [43]. Results from DE analyses are shown in Figures S6-S7 and Table S1, and MEM results are detailed in Figures S7-S8 and Tables S2. Both analyses resulted in FLT profiles largely consistent with each other (Figure S7), as shown before [43]. Our discussion below is primarily based on the results obtained from MEM.

First, for the unmodified DNA (AT10), the donor-only construct (AT10_D) showed a single major lifetime peak (τ_D) at 5.1 ns, which corresponds to the intrinsic lifetime of the donor fluorescence (since there is no acceptor and thus no FRET), consistent with previous results by us and the Wilhelmsson group (Figures 4A & S7A) [43, 50, 67]. In comparison, the DNA containing both the donor and acceptor (AT10_DA) showed a major lifetime peak (τ_{DA}) at 0.27 ns with 86% fractional population with minor peaks at 1.8 ns (7%) and 4.8 ns (7%) (Figures 4A & S7B, Table S2). The major lifetime peak of ~0.3 ns corresponds to a FRET efficiency of ~0.94, which closely matches the calculated FRET of 0.936 for an ideal B-DNA structure [43, 67, 68]. The 4.8 ns lifetime is close to the intrinsic lifetime of the donor in the absence of the acceptor; however, this was not due to an excess of unannealed donor strand, as the same was observed even in the presence of 50% excess acceptor strand (Figure S9). These characteristics of AT10_DA agree well with those of other matched DNA duplexes we had previously examined and confirm that AT10 mainly adopts B-DNA conformation with perhaps a minor population of non-B-DNA conformations [43, 50]. Also, Rad4-binding to the DNA did not alter the FLT profile, as shown previously with other

nonspecific DNA, indicating that nonspecific binding by Rad4 does not lead to detectable changes in tC^{o}-tC_{nitro}-based FRET (Figures 4A & S7C).

In comparison to a single peak profile in the unmodified AT10_D, the donor-only NPOM-modified DNA (AT7_D) showed two peaks: one major peak with a lifetime of 4.5 ns, similar to the τ_D of AT10, but also a minor, 2.0 ns peak (Figures 4B & S7D). This additional 2.0 ns peak was present even for unannealed, single-stranded AT7_D, indicating that it is not sensitive to the DNA's conformation (Figure S7J), and it disappeared upon photocleavage, as seen for AT7_D irradiated for 120 s (AT7_D+$h\nu$), indicating an influence of NPOM on the tC^{o} fluorescence (Figures 4A & S7G). However, despite this minor interference by NPOM on the tC^{o} fluorescence, the donor-acceptor-labeled NPOM-DNA (AT7_DA) showed remarkable resemblance to that of the unmodified DNA (AT10_DA) with one major (0.31 ns (74%)) and two minor peaks (1.8 ns (13%) and 4.8 ns (13 %)) (Figures 4B, 4C & S7E). The similarity between the two DNA constructs indicates that the conformations of NPOM-modified DNA as sensed by the tC^{o}-tC_{nitro} pair are largely unperturbed by the NPOM modification and most retain B-DNA-like conformation. Upon binding to Rad4, however, the FLT profile of AT7_DA changed distinctly compared with unbound DNA, unlike with AT10_DA (Figures 4B, 4D & S7F). Two broader and shorter lifetime peaks (0.16 ns (55%) and 0.39 ns (31%)) replaced the single major peak for unbound AT7_DA at ~0.3 ns while the 1.7 ns and 4.7 ns peaks reduced to 8% and 6% in the fractional population, compared with DNA without Rad4. Such changes in the lifetime distribution translates to an increase in the average FRET efficiency from 0.78 to 0.87 upon Rad4 binding. A broader distribution of lifetimes with multiple peaks in AT7_DA indicated that NPOM-DNA, when specifically bound to Rad4, can access a broader range of distinct conformations with some that deviate from B-DNA. However, the FRET value of the Rad4-bound DNA is significantly different from the FRET E calculated based on the 'open' DNA conformation as seen in the crystal structure of Rad4-bound lesions (0.043), suggesting perhaps a different binding mode for this DNA than other specific substrates [41].

Lastly, NPOM-DNA after photocleavage (AT7+$h\nu$) showed profiles closely resembling that of the unmodified AT10 without or with Rad4, consistent with the expected photoconversion of NPOM-DNA to unmodified DNA (Figure 4A & S7G-I). The small differences in the peak positions and widths were due to the nitrosoacetophenone released after photocleavage, as such differences largely disappeared upon its removal using a G25 size-exclusion resin (Figure S7K). These results reaffirm that light-induced cleavage of the NPOM group from DNA abolishes the specific binding of Rad4 to the DNA while also revealing the unique conformational landscape of NPOM-DNA when it is specifically bound to Rad4.

Progressive, light-induced conversion from specific to nonspecific Rad4-DNA complexes as tracked by FLT.—Seeing that FLT can discern the different conformational landscapes of NPOM-DNA when specifically bound versus nonspecifically bound to Rad4, we next examined progressive changes in FLT of Rad4-bound NPOM-DNA upon gradual increase in photo-irradiation times (0-120 sec) (Figure 2). Progressive increase in photocleavage with increased irradiation time under these conditions resulted in little change in the FLT distributions for free, unbound NPOM-DNA: it mostly retained B-DNA conformation (τ = 0.3 ns) although some broadening of the peaks was observed (Figures

S10A-E). In comparison, the FLT profiles of the Rad4-bound NPOM-DNA changed distinctly with the increased irradiation (Figures 5A & S10F-J). For instance, the two Gaussian peaks at ~0.16 ns and ~0.4 ns had comparable fractional amplitudes (1.9:1) before irradiation but their ratios gradually increased with irradiation (2.8:1 at 60 sec), eventually merging as a single peak with τ of ~0.3 ns, closely resembling the non-specifically bound unmodified AT10_DA (Figure S10J). Interestingly, the same tendency was observed when the ratio between specific and nonspecific binding was altered by progressive change in DNA:protein ratios (Figures 5B & S11-S12). The FLT profiles after 30 sec or 60 sec irradiations resemble the profiles obtained when NPOM-DNA was bound to 2- or 3-fold molar excess of Rad4 (Figure 5C-D). These results indicate that partial irradiation results in a mixture of specifically and nonspecifically bound complexes, as anticipated, yielding conformational distributions that are similar to when there is excess of protein and thus competition between specific and nonspecific binding.

Conformational searches and MD simulations reveal two predominant major groove and one base-displaced intercalated conformation for the NPOM-modified DNA duplex.

Our FLT-FRET study indicates that the NPOM-modified DNA retains a majority B-DNA conformation, at least as sensed by the tC^o and tC_{nitro} FRET probes in these constructs. To gain molecular insights into the FLT-FRET data and understand how the NPOM adduct may impact the DNA duplex structure, we turned to extensive all-atom molecular dynamics (MD) simulations on NPOM-modified DNA. As there is currently no structure available for the NPOM-adduct containing DNA, we first carried out extensive conformational searches to obtain initial models for MD simulations of NPOM-dT in B-DNA (Scheme 1). The search produced five geometry optimized rotamer combinations of NPOM-dT that could fit into the 13-mer B-DNA structure without causing extensive distortions to the duplex (Figure S13). Among these five conformations of NPOM-dT, there were four major groove conformations where NPOM adopted various orientations in the major groove with the dT in *syn* conformation (MJ1, MJ2, MJ3 and MJ4 in Figure S14), and one base-displaced intercalated conformation where NPOM-dT intercalated into the helix with the dT in *anti* conformation and its partner dA extruded into the major groove (INT in Figure S14). We carried out 1.5 μs MD simulations for each of these systems as well as an unmodified control duplex (Figure S15). Among our MD simulations of major groove conformations, one ensemble exhibited denaturing of the duplex and extensive distortions (MJ3 in Figure S15A) and hence was excluded from our further analyses of the structural ensembles.

Our stable 1.5 μs MD simulations of major groove structures (MJ1, 2 and 4) for NPOM-dT converged to two predominant conformations: two rotamers around the long axis of the NPOM rings that placed the nitro group toward the major groove surface (MJ-I) or toward the solvent (MJ-II) (Figures 6 & S15C). These two conformations were observed in all three stable MD ensembles with varying proportions in each population (Figure 6A); they were able to flexibly interchange through different combinations of rotations around the dihedral angles between the NPOM rings and the modified dT (Figures 6 & S16). Of the combined ensembles, 66% adopted either of these two major groove conformations. In the major groove conformation with the nitro group facing the major groove surface (MJ-I; 31% of the population), the NPOM rings were oriented along the helix axis on the major groove surface

with the five-atom ring pointing toward the 5' end of the lesion-containing strand, and its partner dA extruded moderately toward the major groove. With the nitro group facing the solvent (MJ-II; 35% of the population), the NPOM rings were oriented along the base pair planes with the five-atom ring pointing toward its partner dA, protecting the dA from solvent. The remaining 34% of the major groove population were transients that occurred during the transition between the two predominant interchanging rotamers.

For the base-displaced intercalated NPOM-dT, the structural ensemble remained stable, with the NPOM rings intercalated into the DNA helix stacked with neighboring base pairs; the partner dA was flipped into the major groove and protects the NPOM from the solvent (INT-I; Figures 6, S15C & S16). This intercalated conformation comprised 86% of this base-displaced intercalated conformational family. The remaining 14% represented one brief excursion during the MD simulation where the NPOM rings were folded back to stack with dT and stretched the base pair steps (INT-II; Figure S16).

To gain insight on the experimental FLT-FRET data, the two major groove conformations, the intercalated conformation and the unmodified duplex were further analyzed. We modeled the tC^{o}-tC_{nitro} FRET pairs at the respective nucleotide positions and calculated their distances and dihedral angles between the dipoles (detailed in SI Methods). The distances were measured between the center of mass for the middle ring of each fluorophore model (Figure S17A). The dipole dihedral angles were calculated between the modeled dipoles of the FRET pair (Figure S17B). The distance between the FRET pairs was very similar in all conformations of the NPOM-dT-containing DNA and the unmodified DNA: 16.8 ± 1.8 Å for the major groove conformations combined for the two rotamers, 16.7 ± 0.3 Å for the intercalated one, and 16.5 ± 0.1 Å for the unmodified DNA (Figure 6B). While the dipole dihedral angles showed slight differences between the NPOM-dT-containing DNA and the unmodified DNA, the major groove conformations combined for the two rotamers had a value of 170 ± 11° which was close to the unmodified DNA of 182 ± 2°. However, the intercalated conformation was further from the unmodified duplex with a value of 164 ± 5° (Figure 6B). The FRET efficiencies based on the modeled FRET pairs were ~0.96-0.98 for the best representative structures of these conformers, not too far away from the value expected of ideal B-DNA (0.936) and consistent with the FLT experimental results. The major groove conformations also were close to the unmodified DNA in that the Watson-Crick hydrogen bonding was retained at 97-99% occupancy at the two base pairs on each side of the NPOM-dT; however, the *syn* conformation of the NPOM-dT precluded any Watson-Crick base pairing. Hence the distortions induced by the NPOM are very local for the major groove case. On the other hand, the base-displaced intercalated conformation was somewhat more distorted. In addition to the absence of Watson-Crick pairing at the NPOM-dT site, the A:T base pair on the 5' side of the NPOM-dT showed reduced Watson-Crick hydrogen bond occupancies of 81% (N6-H61…O4) and 85% (N3-H3…N1), although the other hydrogen bonds of adjacent base pairs all retained occupancies of 97-99%.

These MD simulations provide atomistic models for the NPOM-DNA and insights into their structural dynamics. While the major groove lesion-containing DNA structures were similar to the unmodified DNA, they also exhibited local lesion conformational dynamics (Figure 6A) that may be relevant to the recognition of the NPOM adduct by Rad4 (see Discussion).

DISCUSSION

Though a variety of photoconvertible DNA/RNA have been shown to modulate DNA/RNA structures and functions, their applications for DNA repair mechanisms have been relatively few. Such chemistry has been used for inducing site-specific, single or double strand breaks in DNA [69-71] and for triggering structural transition in a base excision repair enzyme to study its mechanism [42]. NPOM-DNA was also reported to bind to AlkB, a bacterial direct DNA/RNA repair enzyme that acts on alkylated base damage [72]. NPOM-related, 2-nitrobenzyl or 2-(2-nitrophenyl)propyl groups have also been shown to mask the recognition of a single nucleotide bulge typically recognized by the mismatch repair protein MutS, which photo-irradiation could restore [73].

Nucleotide excision repair (NER) is unique among DNA repair mechanisms in that it repairs an extraordinarily wide range of DNA lesions caused by various environmental and endogenous agents, including intra-strand crosslinks and bulky adduct lesions. A key to its versatility lies in its initial damage sensor protein, Rad4/XPC, that can indirectly detect local thermodynamic destabilization induced by DNA damage without making direct contacts with the lesions themselves. Here in this study, we show for the first time that a photoconvertible modification on DNA can be recognized by Rad4 with specificity similar to that for a *bona fide* lesion (such as the 6-4 photoproduct induced by UV) and that its specific binding can be abolished upon photocleavage of the NPOM-adduct in a light-dose-dependent manner. We also provide, for the first time, valuable structural characterization of the NPOM-modified DNA duplex using FLT-FRET analyses and extensive MD simulations.

Notably, the FLT-FRET studies revealed that NPOM-DNA does not entail a large deviation from the canonical B-DNA form and its specific binding to Rad4 results in an increase in the average FRET. This is in contrast to a previously studied CCC/CCC mismatched DNA which also binds to Rad4/XPC with high specificity. CCC/CCC mismatched DNA, when labeled with tC^{o}-tC_{nitro} FRET probes in analogous positions, showed a broad heterogeneous distributions of fluorescence lifetimes that significantly deviated from B-DNA towards longer lifetime, thus decrease in average FRET [43]. Furthermore, its specific binding to Rad4 further decreased the average FRET, a direction of change in line with expected FRET changes based on the known DNA conformation in the ‘open’ crystal structure [43, 50].These results indicate an intriguing possibility that the conformation of NPOM-DNA when specifically bound to Rad4 may be different from those of CCC/CCC or other DNA lesions that form ‘open’ conformation.

Our subsequent atomistic structures obtained by MD simulation provide novel insights into how NPOM may be recognized by Rad4/XPC as a lesion. The existence of the two nitro group conformations in the major groove is interesting in relation to the recognition by Rad4: the electronegative nitro group, when facing the solvent, may favorably interact with positively charged Arg and Lys amino acids in the DNA binding surface of Rad4 to foster Rad4 binding to a major-groove NPOM conformer. The base-displaced intercalated conformer has a smaller FRET dipole dihedral angle, with somewhat more distorted Watson-Crick pairing, than our unmodified control or the major groove conformers (Figure 6). This smaller dihedral angle represents the well-known intercalation-induced untwisting [74, 75]. In

comparison, untwisting is more modest in the major groove conformers. Base-displaced intercalated conformers have been shown to facilitate Rad4 recognition in computational and experimental work [76, 77]. Computational studies have revealed that the displaced partner base in the case of the NER-proficient (+) *cis*-benzo[*a*]pyrene-dG adduct is readily captured by a pocket between BHD2 and BHD3 while the BHD2 hairpin binds into the minor groove and untwists the duplex; in contrast, these structural hallmarks of initial lesion recognition by Rad4 are missing when the partner nucleotide is absent, in which case the lesion becomes NER-resistant [77]. Different conformers of 2-(acetyl)aminofluorene-dG lesions have also been shown to play a role in their recognition and repair by NER [78, 79]. While it is difficult to ascertain which conformation is prevalent in solution here for NPOM-DNA, base-displaced intercalated conformers can be preferentially represented in specific sequence contexts while being in equilibrium with major groove conformers [80-82]. Further structural studies are needed to reveal the impact of different NPOM-DNA conformations on the recognition by Rad4/XPC and the NER repair.

Overall, our study sets the stage for future studies in which optical triggering can be coupled with a variety of other techniques (e.g., fluorescence conformational dynamics spectroscopy, time-resolved x-ray crystallography, cellular repair kinetic studies, etc.). For instance, NPOM-DNA could be used in monitoring the reverse reaction of the lesion recognition process, using time-resolved crystallography to potentially provide molecular insights into the intermediate steps *en route* to damage recognition. The light-triggered unbinding reaction could also be useful in studying the biochemical processes of NER *in vitro* and *in vivo*, for instance, to probe the timing and structural mechanisms of TFIIH recruitment and lesion verification, albeit in the reverse direction. The large body of literature on protein folding studies shows how deep insights were gained about the folding mechanisms by triggering and monitoring the unfolding of a protein from a folded state. Similarly, triggering an unbinding reaction can provide meaningful insights into the free energy barriers and bottlenecks relevant for the binding reaction; these should be intriguing problems to investigate for the critical steps in biological pathways such as NER *in vitro* and *in vivo*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements.

We thank the members of the Min, Broyde, and Ansari groups.

Funding. This work was funded by National Science Foundation (NSF) grants (MCB-1412692 to J.-H.M and MCB-1715649 to A.A.) and National Institutes of Health grants (R21-ES028384 to J.-H.M, R01-ES025987 to S.B.). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation (NSF) Grant MCB-060037 to S.B., and the high-performance computing resources of New York University (NYU-ITS).

References

1. Greene BI, Hochstrasser RM, Weisman RB and Eaton WA, Proc Natl Acad Sci U S A, 1978, 75, 5255–5259. [PubMed: 281677]

2. Hofrichter J, Sommer JH, Henry ER and Eaton WA, Proc Natl Acad Sci U S A, 1983, 80, 2235–2239. [PubMed: 6572974]
3. Bleger D and Hecht S, Angew Chem Int Ed Engl, 2015, 54, 11338–11349. [PubMed: 26096635]
4. Klan P, Solomek T, Bochet CG, Blanc A, Givens R, Rubina M, Popik V, Kostikov A and Wirz J, Chem Rev, 2013, 113, 119–191. [PubMed: 23256727]
5. Ballister ER, Aonbangkhen C, Mayo AM, Lampson MA and Chenoweth DM, Nat Commun, 2014, 5, 5475. [PubMed: 25400104]
6. Hansen MJ, Feringa FM, Kobauri P, Szymanski W, Medema RH and Feringa BL, J Am Chem Soc, 2018, 140, 13136–13141. [PubMed: 30284823]
7. Schimer J, Pavova M, Anders M, Pachl P, Sacha P, Cigler P, Weber J, Majer P, Rezacova P, Krausslich HG, Muller B and Konvalinka J, Nat Commun, 2015, 6, 6461. [PubMed: 25751579]
8. Xue G, Wang K, Zhou D, Zhong H and Pan Z, J Am Chem Soc, 2019, 141, 18370–18374. [PubMed: 31566962]
9. Liu Y, Zou RS, He S, Nihongaki Y, Li X, Razavi S, Wu B and Ha T, Science, 2020, 368, 1265. [PubMed: 32527834]
10. Gostl R, Senf A and Hecht S, Chem Soc Rev, 2014, 43, 1982–1996. [PubMed: 24413363]
11. Mafy NN, Matsuo K, Hiruma S, Uehara R and Tamaoki N, J Am Chem Soc, 2020, 142, 1763–1767. [PubMed: 31927956]
12. Meldrum RA, Shall S and Wharton CW, Biochem J, 1990, 266, 891–895. [PubMed: 2327973]
13. Liu Q and Deiters A, Acc Chem Res, 2014, 47, 45–55. [PubMed: 23981235]
14. Measey TJ and Gai F, Langmuir, 2012, 28, 12588–12592. [PubMed: 22867440]
15. Salveson PJ, Haerianardakani S, Thuy-Boun A, Kreutzer AG and Nowick JS, J Am Chem Soc, 2018, 140, 5842–5852. [PubMed: 29627987]
16. Courtney TM and Deiters A, Nat Commun, 2019, 10, 4384. [PubMed: 31558717]
17. Grotenbreg GM, Roan NR, Guillen E, Meijers R, Wang JH, Bell GW, Starnbach MN and Ploegh HL, Proc Natl Acad Sci U S A, 2008, 105, 3831–3836. [PubMed: 18245382]
18. Hemphill J, Borchardt EK, Brown K, Asokan A and Deiters A, J Am Chem Soc, 2015, 137, 5642–5645. [PubMed: 25905628]
19. Hemphill J, Chou C, Chin JW and Deiters A, J Am Chem Soc, 2013, 135, 13433–13439. [PubMed: 23931657]
20. Lusic H, Young DD, Lively MO and Deiters A, Org Lett, 2007, 9, 1903–1906. [PubMed: 17447773]
21. Deiters A and Lusic H, Synthesis, 2006, 2006, 2147–2150.
22. Givens RS and Lee J-I, Journal of Photoscience, 2003, 10, 37–48.
23. Luo J, Torres-Kolbus J, Liu J and Deiters A, Chembiochem, 2017, 18, 1442–1447. [PubMed: 28608946]
24. Stroppel AS, Paolillo M, Ziegler T, Feil R and Stafforst T, Chembiochem, 2018, 19, 1312–1318. [PubMed: 29417721]
25. Hanswillemenke A, Kuzdere T, Vogel P, Jekely G and Stafforst T, J Am Chem Soc, 2015, 137, 15875–15881. [PubMed: 26594902]
26. Hemphill J, Liu Q, Uprety R, Samanta S, Tsang M, Juliano RL and Deiters A, J Am Chem Soc, 2015, 137, 3656–3662. [PubMed: 25734836]
27. Berroy P, Viriot ML and Carré MC, Sensors and Actuators B: Chemical, 2001, 74, 186–189.
28. Young DD and Deiters A, Bioorg Med Chem Lett, 2006, 16, 2658–2661. [PubMed: 16513347]
29. Young DD, Lively MO and Deiters A, J Am Chem Soc, 2010, 132, 6183–6193. [PubMed: 20392038]
30. Deiters A, Garner RA, Lusic H, Govan JM, Dush M, Nascone-Yoder NM and Yoder JA, Journal of the American Chemical Society, 2010, 132, 15644–15650. [PubMed: 20961123]
31. Young DD, Govan JM, Lively MO and Deiters A, ChemBioChem, 2009, 10, 1612–1616. [PubMed: 19533711]
32. Govan JM, Uprety R, Hemphill J, Lively MO and Deiters A, ACS Chem Biol, 2012, 7, 1247–1256. [PubMed: 22540192]

33. Young DD, Edwards WF, Lusic H, Lively MO and Deiters A, Chem Commun (Camb), 2008, DOI: 10.1039/b715152g, 462–464. [PubMed: 18188468]
34. Moroz-Omori EV, Satyapertiwi D, Ramel M-C, Høgset H, Sunyovszki IK, Liu Z, Wojciechowski JP, Zhang Y, Grigsby CL, Brito L, Bugeon L, Dallman MJ and Stevens MM, ACS Central Science, 2020, 6, 695–703. [PubMed: 32490186]
35. Zhou W, Brown W, Bardhan A, Delaney M, Ilk AS, Rauen RR, Kahn SI, Tsang M and Deiters A, Angew Chem Int Ed Engl, 2020, 59, 8998–9003. [PubMed: 32160370]
36. Gillet LC and Scharer OD, Chem Rev, 2006, 106, 253–276. [PubMed: 16464005]
37. Marteijn JA, Lans H, Vermeulen W and Hoeijmakers JH, Nat Rev Mol Cell Biol, 2014, 15, 465–481. [PubMed: 24954209]
38. Puumalainen MR, Ruthemann P, Min JH and Naegeli H, Cell Mol Life Sci, 2016, 73, 547–566. [PubMed: 26521083]
39. Kraemer KH, Patronas NJ, Schiffmann R, Brooks BP, Tamura D and DiGiovanna JJ, Neuroscience, 2007, 145, 1388–1396. [PubMed: 17276014]
40. Paul D, Mu H, Zhao H, Ouerfelli O, Jeffrey PD, Broyde S and Min JH, Nucleic Acids Res, 2019, 47, 6015–6028. [PubMed: 31106376]
41. Min JH and Pavletich NP, Nature, 2007, 449, 570–575. [PubMed: 17882165]
42. Lee S, Radom CT and Verdine GL, Journal of the American Chemical Society, 2008, 130, 7784–7785. [PubMed: 18507380]
43. Chakraborty S, Steinbach PJ, Paul D, Mu H, Broyde S, Min JH and Ansari A, Nucleic Acids Res, 2018, 46, 1240–1255. [PubMed: 29267981]
44. Borjesson K, Preus S, El-Sagheer AH, Brown T, Albinsson B and Wilhelmsson LM, J Am Chem Soc, 2009, 131, 4288–4293. [PubMed: 19317504]
45. Preus S, Borjesson K, Kilsa K, Albinsson B and Wilhelmsson LM, J Phys Chem B, 2010, 114, 1050–1056. [PubMed: 20039634]
46. Dumat B, Larsen AF and Wilhelmsson LM, Nucleic Acids Res, 2016, 44, e101. [PubMed: 26896804]
47. Chen X, Velmurugu Y, Zheng G, Park B, Shim Y, Kim Y, Liu L, Van Houten B, He C, Ansari A and Min JH, Nat Commun, 2015, 6, 5849. [PubMed: 25562780]
48. Cuellar RE, Ford GA, Briggs WR and Thompson WF, Proc Natl Acad Sci U S A, 1978, 75, 6026–6030. [PubMed: 366608]
49. Palais R and Wittwer CT, in Computer Methods, Part A, Academic Press, 2009, vol. 454, pp. 323–343.
50. Paul D, Mu H, Tavakoli A, Dai Q, Chen X, Chakraborty S, He C, Ansari A, Broyde S and Min J-H, Nucleic Acids Research, 2020, 48, 12348–12364. [PubMed: 33119737]
51. Velmurugu Y, Chen X, Slogoff Sevilla P, Min JH and Ansari A, Proc Natl Acad Sci U S A, 2016, 113, E2296–2305. [PubMed: 27035942]
52. Livesey AK and Brochon JC, Biophys J, 1987, 52, 693–706. [PubMed: 19431708]
53. Steinbach PJ, Anal Biochem, 2012, 427, 102–105. [PubMed: 22504734]
54. Steinbach PJ, Ionescu R and Matthews CR, Biophysical Journal, 2002, 82, 2244–2255. [PubMed: 11916879]
55. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich A, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JAJ, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB and Fox DJ, Gaussian 09 Revision E.01, Gaussian Inc.: Wallingford CT, 2016.
56. Case DA, Ben-Shalom IY, Beozell SR, Cerutti DS, Cheatham I, T.E.,, Cruzeiro VWD, Darden TA, Duke RE, Ghoreishi D, Gilson MK, Gohlke H, Goetz AW, Greene D, Harris R, Homeyer N,

Huang Y, Izadi S, Kovalenko A, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Mermelstein DJ, Merz KM, Miao Y, Monard G, Nguyen C, Nguyen H, Omelyan I, Onufriev A, Pan F, Qi R, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Smith J, Salomon-Ferrer R, Swails J, Walker RC, Wang J, Wei H, Wolf RM, Wu X, Xiao L and York DM, AMBER 2018, University of California, San Francisco: CA, 2018.

57. McGall GH, Barone AD, Diggelmann M, Fodor SPA, Gentalen E and Ngo N, Journal of the American Chemical Society, 1997, 119, 5081–5090.
58. Boháčová S, Vaníková Z, Poštová Slavětínská L and Hocek M, Organic & Biomolecular Chemistry, 2018, 16, 5427–5432. [PubMed: 29905748]
59. Ovcharenko A, Weissenboeck FP and Rentmeister A, Angewandte Chemie International Edition, 2020, n/a.
60. Vaníková Z, Janoušková M, Kambová M, Krásný L and Hocek M, Chemical Science, 2019, 10, 3937–3942. [PubMed: 31015933]
61. Liu M, Jiang S, Loza O, Fahmi NE, Sulc P and Stephanopoulos N, Angew Chem Int Ed Engl, 2018, 57, 9341–9345. [PubMed: 29790232]
62. Scharer OD, Cold Spring Harb Perspect Biol, 2013, 5, a012609. [PubMed: 24086042]
63. Shi Y, Dierckx A, Wanrooij PH, Wanrooij S, Larsson NG, Wilhelmsson LM, Falkenberg M and Gustafsson CM, Proc Natl Acad Sci U S A, 2012, 109, 16510–16515. [PubMed: 23012404]
64. Xia S, Wood M, Bradley MJ, De La Cruz EM and Konigsberg WH, Nucleic Acids Res, 2013, 41, 9077–9089. [PubMed: 23921641]
65. Posse V, Hoberg E, Dierckx A, Shahzad S, Koolmeister C, Larsson NG, Wilhelmsson LM, Hallberg BM and Gustafsson CM, Nucleic Acids Res, 2014, 42, 3638–3647. [PubMed: 24445803]
66. Klostermeier D and Millar DP, Biopolymers, 2001, 61, 159–179. [PubMed: 11987179]
67. Sandin P, Borjesson K, Li H, Martensson J, Brown T, Wilhelmsson LM and Albinsson B, Nucleic Acids Res, 2008, 36, 157–167. [PubMed: 18003656]
68. Preus S, Kilsa K, Miannay FA, Albinsson B and Wilhelmsson LM, Nucleic Acids Res, 2013, 41, e18. [PubMed: 22977181]
69. Dussy A, Meyer C, Quennet E, Bickle TA, Giese B and Marx A, Chembiochem, 2002, 3, 54–60. [PubMed: 17590954]
70. Ordoukhanian P and Taylor JS, Bioconjug Chem, 2000, 11, 94–103. [PubMed: 10639091]
71. Zhang K and Taylor JS, Biochemistry, 2001, 40, 153–159. [PubMed: 11141065]
72. Yi C and He C, SCIENCE CHINA Chemistry, 2009, 53, 86–90.
73. Seio K, Ohno Y, Ohno K, Takeshita L, Kanamori T, Masaki Y and Sekine M, Bioorg Med Chem Lett, 2016, 26, 4861–4863. [PubMed: 27567370]
74. Jain SC and Sobell HM, J Biomol Struct Dyn, 1984, 1, 1179–1194. [PubMed: 6400817]
75. Dikic J and Seidel R, Biophys J, 2019, 116, 1394–1405. [PubMed: 30954211]
76. Mu H, Zhang Y, Geacintov NE and Broyde S, Chem Res Toxicol, 2018, 31, 1260–1268. [PubMed: 30284444]
77. Mu H, Geacintov NE, Min JH, Zhang Y and Broyde S, Chem Res Toxicol, 2017, 30, 1344–1354. [PubMed: 28460163]
78. Yeo JE, Khoo A, Fagbemi AF and Scharer OD, Chem Res Toxicol, 2012, 25, 2462–2468. [PubMed: 23088760]
79. Jain V, Hilton B, Patnaik S, Zou Y, Chiarelli MP and Cho BP, Nucleic Acids Res, 2012, 40, 3939–3951. [PubMed: 22241773]
80. Jain V, Vaidyanathan VG, Patnaik S, Gopal S and Cho BP, Biochemistry, 2014, 53, 4059–4071. [PubMed: 24915610]
81. Stavros KM, Hawkins EK, Rizzo CJ and Stone MP, Nucleic Acids Res, 2014, 42, 3450–3463. [PubMed: 24366876]
82. Wang F, Elmquist CE, Stover JS, Rizzo CJ and Stone MP, Biochemistry, 2007, 46, 8498–8516. [PubMed: 17602664]

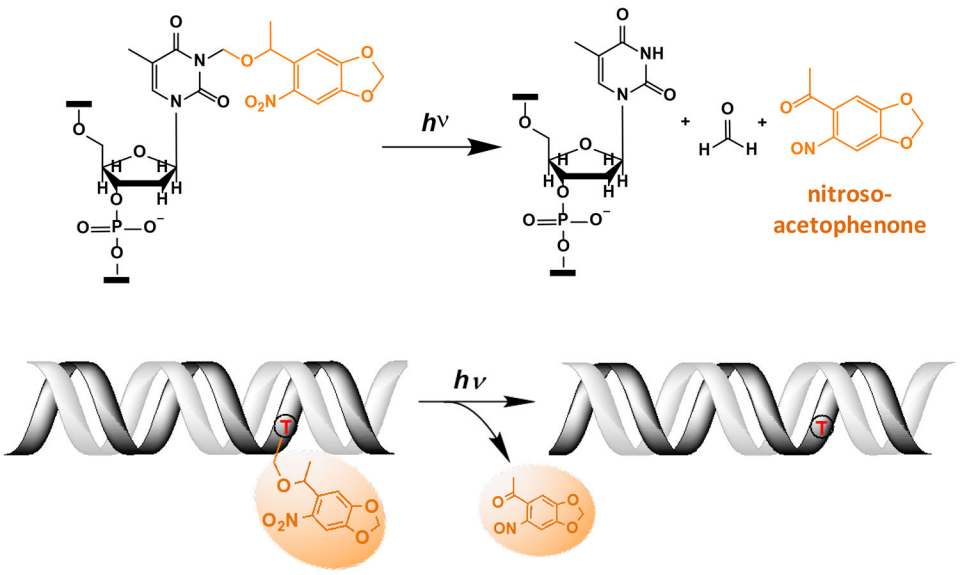


Figure 1. Light-induced photocleavage reaction of the NPOM group from DNA.

(*Top*) Schematic of the photocleavage reaction. Upon light irradiation (λ=365 nm or 405 nm), the NPOM group (orange) is cleaved from the modified thymidine (dT) in the DNA, restoring unmodified thymidine while releasing nitrosoacetophenone. (*Bottom*) Cartoon of the photocleavage reaction in NPOM-containing duplexed DNA.

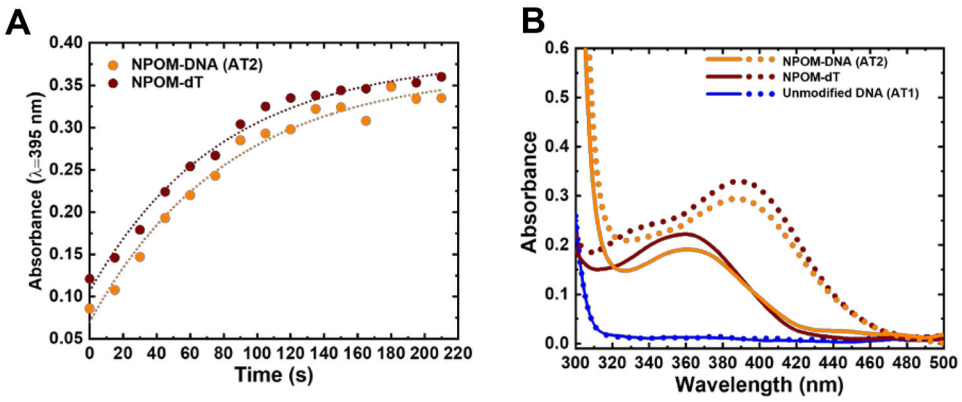
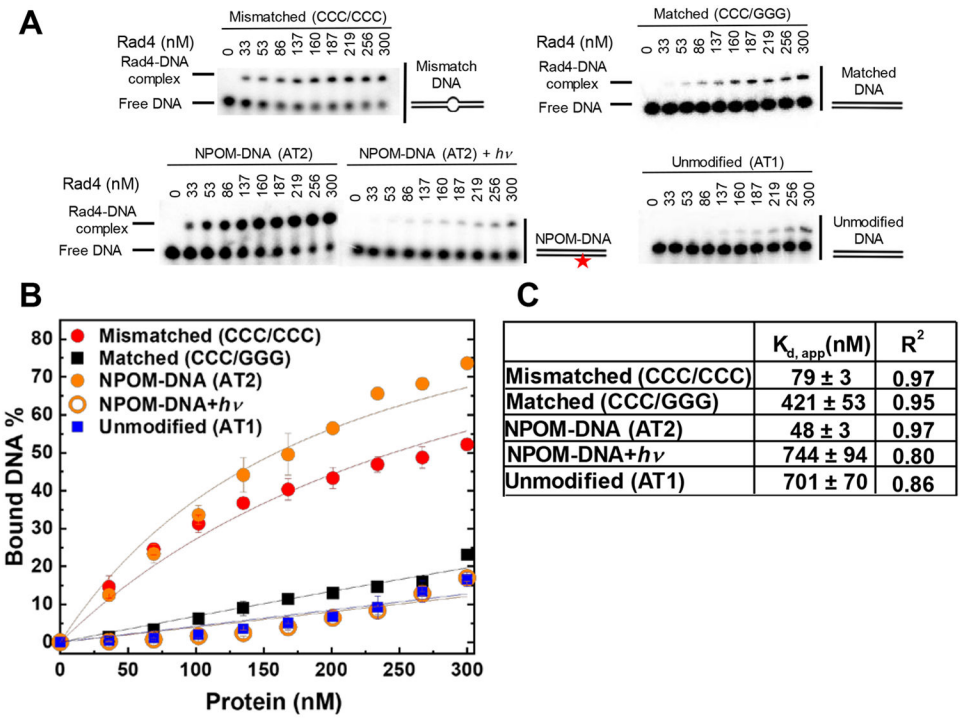


Figure 2. Dose-dependent photocleavage of NPOM from DNA as monitored *in situ* by UV-visible absorption spectroscopy.

(A) Absorption at 395 nm versus time for NPOM-DNA (AT2, orange) and NPOM-dT nucleoside (brown). Dotted lines indicate single exponential fitting of the data (see also Figures S1 and S2). **(B)** Absorption spectra of NPOM-DNA (AT2, orange), NPOM-dT (brown), and unmodified DNA duplex (AT1, blue) before (solid line) and after 120 s of light irradiation (dotted line).



	$K_{d,app}$(nM)	R^2
Mismatched (CCC/CCC)	79 ± 3	0.97
Matched (CCC/GGG)	421 ± 53	0.95
NPOM-DNA (AT2)	48 ± 3	0.97
NPOM-DNA+*hν*	744 ± 94	0.80
Unmodified (AT1)	701 ± 70	0.86

Figure 3. Apparent Rad4-binding affinities of DNA constructs measured by competition electrophoretic mobility shift assays (EMSA).

(A) Typical gel images showing the wild-type Rad4-Rad23 complex binding to various DNA constructs. "NPOM-DNA + hν" indicates NPOM-DNA photocleaved by λ=365 nm light applied for 3 min. Mismatched (CCC/CCC) and matched (CCC/GGG) DNA represent typical specific and nonspecific binding substrates, respectively. The sequences of DNA are in Table S1. **(B)** Quantification of the Rad4-bound DNA fractions versus total concentrations of the protein from gels including those shown in (B). The symbols and error bars indicate the means and ranges as calculated by ± sample standard deviations, respectively, from triplicate experiments. Solid lines indicate the fit curves of the data point. **(C)** $K_{d,app}$ and R^2 of the fits derived from (B).The errors indicate the errors of the nonlinear regression fit.

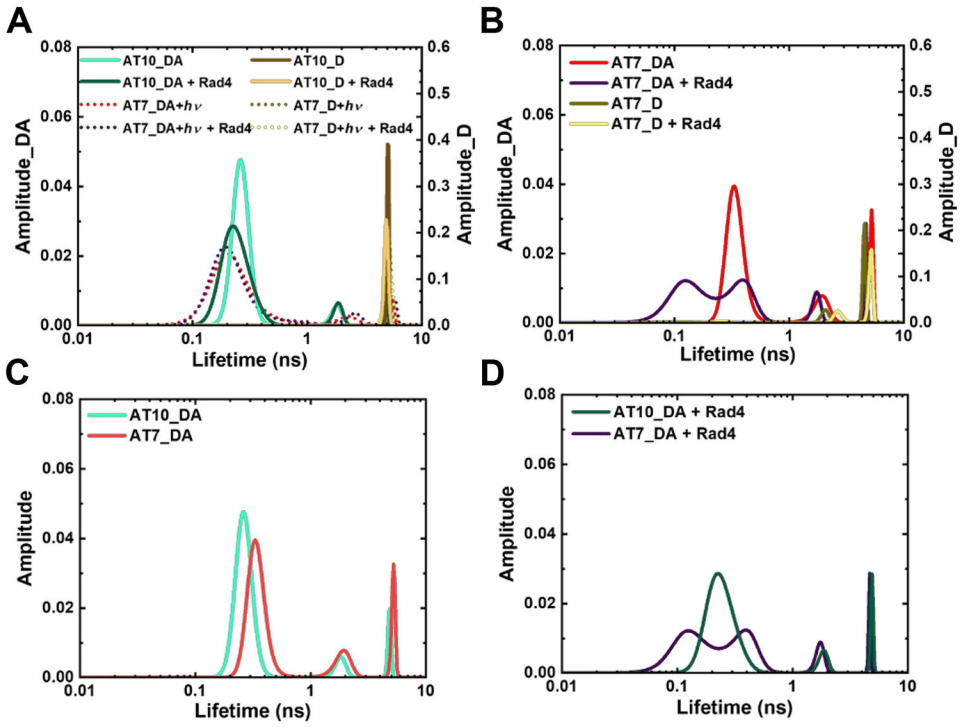


Figure 4. Fluorescence lifetime distributions obtained from MEM analyses for various tC^{o}-tC_{nitro}-labeled DNA and DNA-protein complexes.

"_D" indicate DNA with donor only; "_DA" indicate DNA with donor/acceptor pair. **(A)** Unmodified DNA (**AT10**) in the absence and presence of Rad4 and its comparison with NPOM-DNA after 120 s of photocleavage reaction (**AT7+hν**). **(B)** NPOM-modified DNA (**AT7**) in the absence and presence of Rad4. **(C, D)** Overlay of unmodified (AT10_DA) and NPOM-DNA (AT7_DA) without Rad4 **(C)** and in the presence of Rad4 **(D)** Reproducibility of MEM FLT distributions for each DA sample is shown in Figure S8. Full reports of the lifetimes, fractional amplitudes, FRET efficiencies of each peak as well as the sample's average FRET efficiencies are in Table S2.

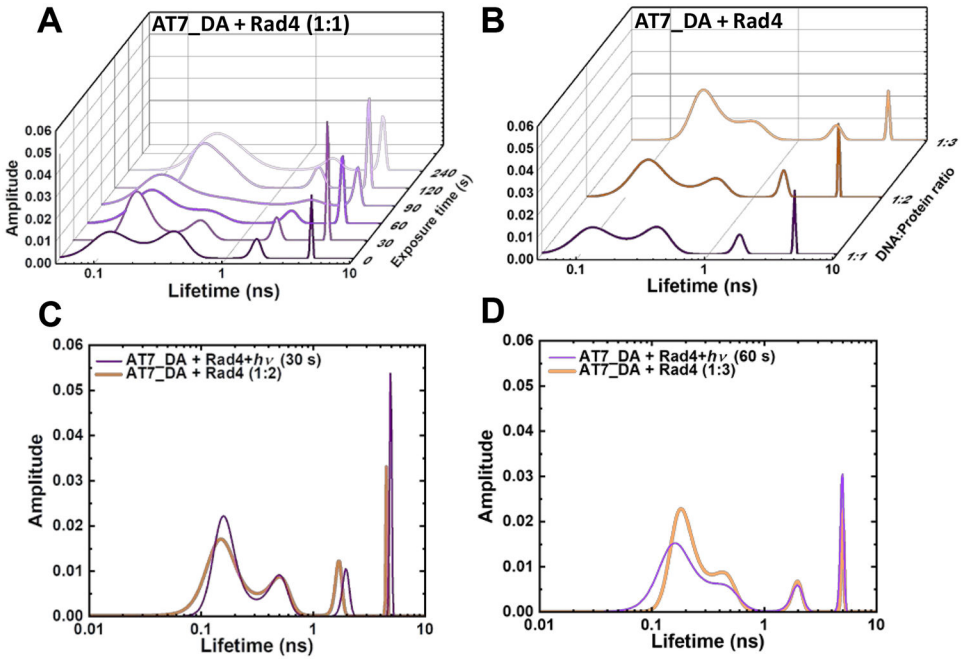


Figure 5. Light-induced conversion from specific to nonspecific Rad4-DNA complexes as tracked by FLT.

(A) FLT distributions obtained from MEM analyses of Rad4-bound NPOM-DNA (AT7_DA) irradiated with varying photocleavage times (0-240 s). **(B**) FLT distributions of NPOM-DNA with varying DNA:Rad4 ratios. **(C)** Overlay of 30 s irradiation and 1:2 AT7_DA:Rad4 complex. **(D)** Overlay of 60 s irradiation and 1:3 AT7_DA:Rad4 complex.

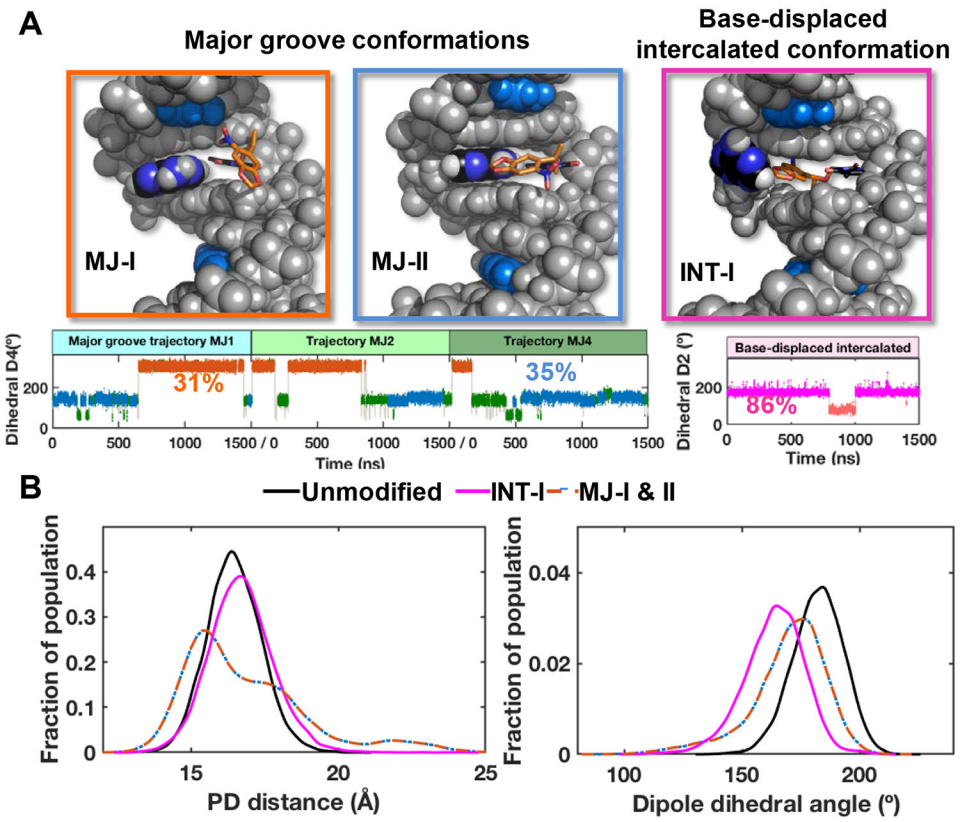
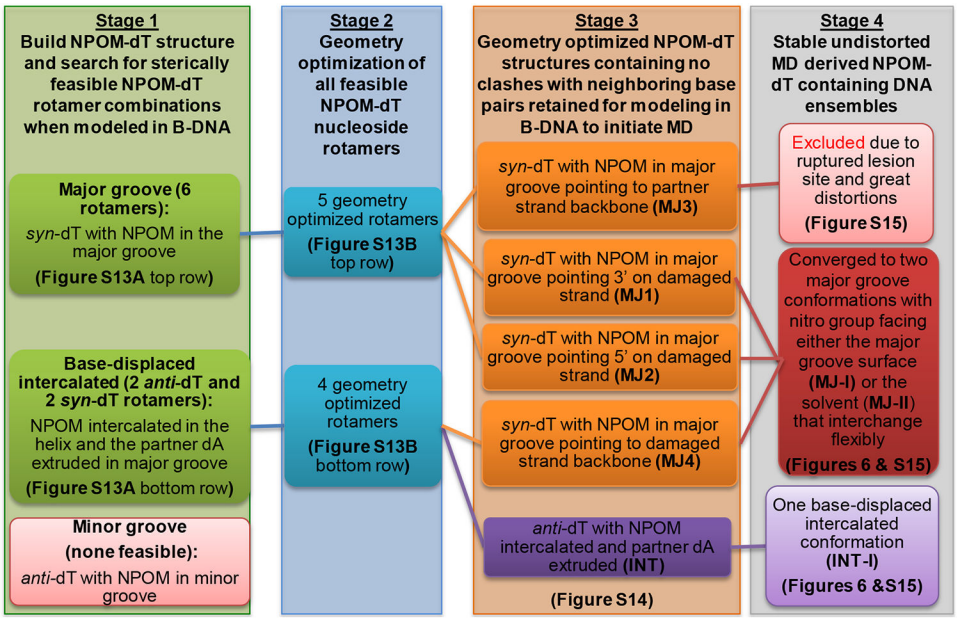


Figure 6. NPOM-containing DNA structures obtained from stable MD derived ensembles.

(A) Best representative structures for the two major groove (MJ) and one base-displaced intercalated (INT) conformations. The NPOM-dT dihedral angles (Figure S16) that determine these different conformations are shown for each conformational family (MJ and INT) with each cluster color-coded and labeled with its percentage of population. The transient clusters in the major groove ensembles are colored green. These structures are shown in cartoon in Figure S15. **(B)** The distributions of the modeled FRET pair distances (PD distance) and dipole dihedral angles for the major groove and base-displaced intercalated NPOM-dT-containing DNA and unmodified DNA. The definitions for the PD distance and dipole dihedral angle are given in Figure S17. The major groove values are for the two dominant major groove conformations (MJ-I and MJ-II, 66% of the population), and the base-displaced intercalated values are for the stable intercalated conformation (INT-I, 86% of the population).



Scheme 1.
Conformational search strategy for NPOM-dT-containing DNA

Table 1.

Sequences of the DNA duplexes used in this study

DNA	Sequences	T_m (°C)
Mismatched (**CCC/CCC,** CH10_NX)	5′-TTGACTCGACATCC**CCC**GCTACAA -3′ 3′- ACTGAGCTGTAGG**CCC**CGATGTTA -	61.0 ± 0.8[c]
Matched (**CCC/GGG,** CH7_NX)	5′-TTGACTCGACATCC**CCC**GCTACAA -3′ 3′- ACTGAGCTGTAGG**GGG**CGATGTTA -	76.5 ± 1.0[c]
NPOM-DNA (AT2)	5′-TTGACTCGACATCC**GAA**GCTACAA -3′ 3′- ACTGAGCTGTAGG**CTT**CGATGTTA -	45.2 ± 0.2
Unmodified (AT1)	5′-TTGACTCGACATCC**GAA**GCTACAA -3′ 3′- ACTGAGCTGTAGG**CTT**CGATGTTA -	52.0 ± 0.0
NPOM-DNA [a, b] (AT7_DA)	5′-TTGACTCGACATC**P****GAA**GGTACAA -3′ 3′- ACTGAGCTGTAGG**CTT**C**D**ATGTTA -	47.1 ± 0.1
Unmodified-DNA[a] (AT10_DA)	5′-TTGACTCGACATC**P****GAA**GGTACAA -3′ 3′- ACTGAGCTGTAGG**CTT**C**D**ATGTTA -	53.6 ± 0.5

[a] D: tC^O(FRET donor); P:tC_{nitro}(FRET acceptor). This design is based on design in ref. 43

[b] Red indicates NPOM-modified dT.

[c] These values are from ref. 40.

Bold indicates the position of the 3-bp sequence corresponding to the CCC/CCC mismatched site in CH10_NX.

T_m values are reported as the average ± standard deviation (s.d.) of three independent measurements (see Figure S3). The minimal uncertainty in T_m as judged by half the temperature interval between successive data points in the derivatives graph is 0.5 °C.