Single Molecule Studies of DNA Damage Recognition by XPA

by

Emily Celeste Beckwitt

B.A., Colby College, 2010

Submitted to the Graduate Faculty of the

School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Emily Celeste Beckwitt

It was defended on

November 8, 2019

and approved by

Jeffrey Brodsky, Professor, Department of Biological Sciences

Marcel Bruchez, Professor, Departments of Biological Sciences and Chemistry, Carnegie Mellon University

Patricia Opresko, Professor, Department of Environmental and Occupational Health

Hong Wang, Associate Professor, Department of Physics, North Carolina State University

Dissertation Director: Bennett Van Houten, Professor, Department of Pharmacology and Chemical Biology

Single Molecule Studies of DNA Damage Recognition by XPA

Emily Celeste Beckwitt, PhD

University of Pittsburgh, 2019

Nucleotide excision repair (NER) is responsible for the repair of a wide range of DNA lesions, including UV-induced photoproducts and bulky base adducts. XPA is an essential protein in eukaryotic NER, although questions about its stoichiometry and mechanism of damage recognition have been heretofore unresolved. Regions of intrinsic disorder within the N- and C-termini of XPA have made structural work on the full-length protein challenging and compel an alternative approach. We have used PeakForce Tapping® atomic force microscopy to show that human XPA binds to DNA as a monomer and bends it ~60°. Furthermore, XPA demonstrated specificity for the helix-distorting base adduct, N-(2'-deoxyguanosin-8-yl)-2-acetylaminofluorene. Single molecule fluorescence microscopy revealed that DNA-bound XPA exhibits multiple modes of linear diffusion between paused phases. These included long distance motion with rapid diffusion ($D \approx 0.04\ \mu m^2/s$) consistent with hopping and short distance motion ($D \approx 0.0003\ \mu m^2/s$) consistent with sliding along the DNA contour. The presence of DNA damage increases pausing by proteins undergoing one-dimensional target search. A truncated mutant, lacking most of the intrinsically disordered regions and made up of just residues 98-239 of the DNA binding domain, exhibits less pausing on UV-damaged DNA compared to the full length protein. In summary, our data are consistent with a model in which the conformational state of XPA is dependent upon the presence of DNA damage and bending.

Table of Contents

List of Tables

List of Figures

List of Equations

Preface

As Sandra Day O'Connor once said, and Leslie Knope of *Parks and Recreation* later reminded us, "We don't accomplish anything in this world alone." Reflecting on the journey I have taken during my graduate work, these words feel especially true. First thanks must go to my advisor, Ben Van Houten. Not only did he give me the opportunity to begin work on this project, I am certain that I would not have finished without his consistent support and guidance. Ben's genuine enthusiasm for scientific research is a driving force in the lab, pushing us all to do our best work.

I also have immense gratitude for all lab members who have helped me throughout the last five years. Muwen Kong trained me on all single molecule techniques when I joined the lab. His patience, knowledge, and companionship were invaluable to me and so appreciated. Thanks must also go to all undergraduate and rotation students who I have worked with over the course of this project, including Samuel Johnson, Nacef Guess, Alex Bowman, and Halima Alnaqbi. Isadora Carnaval Detweiler's contagious good attitude and persistent efforts directly contributed to this dissertation. Thank you all for trusting in my guidance and giving your time to this lab. Special thanks must go to Sunbok Jang, whose biochemical expertise and willingness to teach me and help with this project have been crucial to its success, and to Namrata Kumar and Katie Lemon for their friendship and always offering to bring me a snack when I was too stressed to eat. Other lab members, including Maria Beecher, Vera Roginskaya, Wei Qian, and Lili Liu have all also generously offered both moral and scientific support throughout this process.

I must also thank my committee, who has provided thoughtful feedback and advice at all stages of this dissertation. Thanks to Patty Opresko, Jeff Brodsky, Marcel Bruchez, and Hong

Wang. I am grateful for your wealth of knowledge and that you have helped to steer me and this project in the right direction. Thanks must also go to Simon Watkins for his involvement and to him and Callen Wallace for their help with the single molecule fluorescence microscopy. I am also, of course, grateful to the Molecular Biophysics and Structural Biology Program at the University of Pittsburgh and Carnegie Mellon University, both students and faculty. Collaborations with the labs of Caroline Kisker (including Jochen Kuper and Florian Sauer) at the University of Würzburg and Thomas Carell (including Nina Simon and Johanna Bretzler) at the University of Munich have been invaluable to the XPA project.

I must thank my dear friends and colleagues, who I met during first year classes, before we all dispersed and joined different labs: Tiffany Bernardo, Ryan Staudt, Dillon Kunkle, and Josh Lorenz-Guertin. You have truly made grad school a joy and I will cherish our friendships forever. Finally, thanks to my family. My non-blood related sisters, Aneth Canale and Kristen Devlin, have been with me through undergraduate and graduate school, and all of life in between. Mom, Dad, and Kate, I know I would not have completed this dissertation without your support. Thank you for always encouraging me to trust myself and for making a home where I feel safe and warm. I love you and I thank you.

To all who have helped me successfully complete this dissertation—all names listed above and those not included here—I dedicate this to you.

1.0 Introduction

1.1 Protein Target Search on DNA

In order for any DNA-binding protein to function, it must find its target. This is true for proteins like transcription factors and restriction enzymes that bind specific DNA sequences as well as DNA repair proteins that bind to specific forms of base damage. In a human cell, which contains a high ratio of non-specific/specific DNA binding sites, this can be seen as the biological equivalent of searching for a needle in a haystack of 10^9 base pairs.

Early work by von Hippel and colleagues have laid the foundation for subsequent studies on protein-DNA target search. An important series of their papers laid out theoretical calculations relating to diffusion and applied them to *lac* repressor, a bacterial protein which regulates gene expression by binding to the *lac* operon sequence in DNA[1-4]. Free proteins in solution undergo Brownian motion and must rely on random collision with the appropriate site to find their target. Furthermore, the properties required for recognition drop off significantly if the protein binds even one base pair (0.34 nm) away from the target[5]. Still, three-dimensional diffusion may be an effective method if protein concentrations are sufficiently high, thereby increasing the probability of a specific collision[6]. However, many DNA-binding proteins are expressed at relatively low levels. For example, *E. coli* cells only contain about ten copies of the *lac* repressor[7]. Similarly, one HeLa cell contains between 2.5 and 8×10^4 molecules of XPC, a protein involved in damage recognition during nucleotide excision repair[8,9]; this is several orders of magnitude less than the number of possible non-specific binding sites. In cases like these, a three-dimensional search may

actually take longer than a one-dimensional search[i]. By reducing target search to just one dimension, a relatively low copy number of search proteins can find their target more efficiently than via a purely three-dimensional process[10,11].

Biological evidence for this type of search strategy was first reported for the *lac* repressor, which is able to find the *lac* operon sequence faster than theoretically possible for a three-dimensional search[11]. Consequently, this process has been deemed facilitated diffusion[1,2,4]. Four major categories of facilitated diffusion have been defined: sliding, hopping (microscopic dissociation-reassociation events), jumping (macroscopic dissociation-reassociation events), and intersegmental transfer between segments within the same DNA molecule[1]. Typically, it is understood that the protein will bind DNA at any site, due to some affinity for non-specific DNA, and then transition from a three-dimensional to a one-dimensional search, undergoing some combination of one or more facilitated diffusion processes. This two-step reaction scheme has been suggested by early theoretical work for molecular interactions in general[10,12], and protein-DNA binding in particular[1].

Later work by Slutsky, Mirny, and others, have investigated the role of protein-DNA energy binding landscapes in recognition of and binding to specific DNA sites. A protein undergoing a random walk during linear diffusion encounters a wide variety of energy potentials for DNA-binding[13]. Slutsky and Mirny proposed a model in which proteins must diffuse rapidly along vast sequences of DNA in order to adequately sample the DNA. This type of diffusion requires a relatively smooth binding energy landscape and that proteins are interacting with DNA

[i] Please refer to Section 2.6.5 for further discussion of the limits of one-dimensional diffusion, in the context of a protein translocating along DNA *in vitro*.

relatively loosely. The authors conclude such a mechanism dictates that proteins encounter and overcome relatively small free energy barriers, less than 1–2 × k_BT (where k_BT is the thermal energy term: the product of the Boltzmann constant, k_B, and temperature, T, or ~4.11 × 10^{-21} J at 25°C)[5,14]. The smooth energy landscape prevents the protein from getting trapped in any one position and allows for acceleration of target search. However, the interactions required to recognize a target site and form a stable protein-DNA complex require an energy landscape rugged enough to inhibit protein diffusion. This can only be achieved with free energy barriers greater than 5 × k_BT (refs. [5,14]). This concept has been named the search-speed/stability paradox, as these two states have mutually exclusive energy requirements[5]. To reconcile these conflicting requirements for target site recognition, Slutsky and Mirny proposed a two-state model, whereby the protein can adopt two conformations[5,14]. The "search state" corresponds to a structural conformation allowing for fast linear diffusion and smooth energy landscapes with standard deviation $\sigma \lesssim 1–2 \times k_BT$. The "recognition state" corresponds to a rugged energy landscape, $\sigma \gtrsim 5 \times k_BT$, and higher affinity complexes. Hu and Shklovskii also report that energetic disorder, leading to rugged energy landscapes, slows linear diffusion of proteins bound to DNA[15].

Single molecule studies by van Oijen and colleagues provide compelling evidence that the transcription factor p53 adopts these two conformations (i.e. a search state and a recognition state) during search for its cognate sequence[16,17]. The authors calculated diffusion constants from single particle tracking data and used this to predict and interpret corresponding energy landscapes. Their results indicate that p53 switches between two conformations: a search state with major contacts between C-terminal domains and DNA, and a recognition state where the core domains fold in, providing additional contacts with the DNA and resulting in slower diffusion[16]. These principles

have also been studied for DNA repair proteins, including Msh2-Msh6[18], thymine DNA glycosylase[19], MutS[20], UV-DDB[21], Rad4[22], and PARP1[23].

1.2 Nucleotide Excision Repair

Our genomes are subject to constant assault and suffer approximately 10,000 to 70,000 lesions per cell per day[24]. Nucleotide excision repair (NER) is a highly conserved DNA repair pathway that is able to specifically recognize and repair a wide range of structurally and chemically distinct DNA lesions. In humans, this process involves approximately 30 proteins, working together to protect our genomes from the damaging effects of UV radiation and chemical carcinogens.

1.2.1 DNA Lesions Repaired by NER

Though diverse, the majority of lesions repaired via the NER pathway destabilize or distort the DNA helix in some way. UV radiation (254 nm) causes formation of two major lesions in DNA: the cyclobutane pyrimidine dimer (CPD) and 6-4 photoproduct (6-4PP), at a ratio of approximately three to one[25,26]. NMR studies have shown that 6-4PP lesions cause significant bending (44°) in the DNA helix and disrupt hydrogen bonding; in contrast, DNA with a CPD lesion maintains a B-form helix with a 9° bend[27]. Cisplatin, commonly used in cancer chemotherapy, readily forms covalent attachments to purines in DNA, resulting in intrastrand crosslinks, interstrand crosslinks, and monoadducts[28]. The coordination complex containing cisplatin may be reversible with a strong reductant, such as cyanide. These adducts, and other

platinum derivatives, can be repaired via NER[29]. Furthermore, a diverse group of polycyclic aromatic hydrocarbons and aromatic amines are also recognized and repaired by NER. Acetylaminofluorene (AAF), a synthetic carcinogen which produces lesions at the C8 position of guanine, is one such well-studied substrate[30-33].

1.2.2 Steps in Eukaryotic NER

The general steps of eukaryotic NER are illustrated in Figure 1. Like all DNA repair pathways, NER begins with a damage detection step. The NER damage detection step can be initiated in two general ways: during transcription or in chromatin[34-36]. During transcription-coupled (TC) NER, RNA polymerase stalls at a site of damage and is recognized by Cockayne syndrome protein A (CSA) and Cockayne syndrome protein B (CSB), which promote removal of the polymerase from the damage site and recruitment of subsequent repair proteins. Recently, broader roles for CSA and CSB in proteasome-mediated degradation of an immediate early gene product and transcription restart after UV have been reported[37].

Global genome (GG) NER is initiated by UV-damaged DNA binding protein (UV-DDB) and/or XPC-RAD23B-CETN2 when they recognize the lesion at any site in the chromatin. UV-DDB, a major sensor of UV photoproducts during GG NER, exists as a heterodimer of subunits DDB1 and DDB2. It is associated with the CUL4A-RBX E3 ubiquitin ligase which modifies core histones in response to UV radiation[38-40]. Single molecule analysis of the dynamics of UV-DDB binding to damaged DNA have indicated a conformational proofreading mechanism, where binding of UV-DDB at a site of DNA damage induces a conformational change in the protein which stabilizes the UV-DDB-lesion complex[21]. Rad4-Rad23 (yeast homolog of XPC-RAD23B) has been shown to exhibit anomalous subdiffusion during recognition of CPD lesions in DNA[22].

This “recognition at a distance” allows for the assembly of subsequent repair events without steric interference. As discussed in the previous section, the NER recognition proteins appear to have specificity for DNA with disrupted base pairing or a thermodynamically destabilized helix, even in the absence of a traditional lesion[36,41,42].

The two NER pathways converge for a damage verification step, where the TFIIH helicase complex unwinds the DNA and tests for the presence of damage. The seven core subunits of TFIIH are essential for both transcription and NER: XPB, XPD, p62, p52, p44, p34, and p8/TTDA[43]. An additional CDK-activating kinase domain is required for transcription, but not for NER. Replication protein A (RPA) binds non-damaged single stranded DNA, stabilizing the pre-incision complex. After damage verification, endonucleases XPF-ERCC1 and XPG make sequential incisions, 5’ and 3’ of the lesion respectively. After incision by XPF-ERCC1, and before the XPG nuclease is activated, DNA polymerase (δ, ε, or κ) and PCNA assemble to begin DNA synthesis to fill in the gap. Then incision by XPG releases the lesion-containing 24-32 base oligonucleotide[44,45]. Finally, the new DNA backbone is sealed by DNA ligase (I or III)[36].

The protein XPA, discussed in depth in subsequent sections, is an essential protein in both TC and GG NER. It has no known enzymatic function. Initial reports on the function of XPA in NER conclude that it is involved in early steps of damage recognition[46], although more recent models place XPA later in the pathway[47], acting as a scaffold and interacting with other NER proteins[48-54], and RPA in particular[55-57]. Though indispensable, the precise role of XPA during NER remains unclear, and likely is dependent on multiple factors, including type of damage and presence/absence of other factors. Importantly, it has been suggested that damage recognition in NER, which needs to be highly specific to a diverse range of structures, is accomplished via a “discrimination cascade” involving multiple proteins, each with imperfect selectivity[58-60]. In

support of this model, XPA enhances the damage specificity of TFIIH by promoting both its translocation along non-damaged DNA and stalling at a lesion[61,62]. As such, it remains of significant interest to investigate how XPA interacts with DNA lesions.

1.2.3 Diseases Associated with NER

Genetic mutations affecting NER proteins can cause multiple autosomal recessive disorders, reviewed in ref.[63]. These diseases are rare but come with significant challenges and decreased life expectancies. One major genetic disorder associated with defective NER is xeroderma pigmentosum (XP). There are seven types of XP, defined by complementation groups and named for the specific gene carrying the mutation: *XPA*, *ERCC3* (XPB), *XPC*, *ERCC2* (XPD), *DDB2* (XPE), *ERCC4* (XPF), and *ERCC5* (XPG). An eighth type of XP, named XP variant (XPV), is caused by mutations in the gene coding for DNA polymerase η, involved in translesion DNA synthesis and not NER. XP patients suffer from extreme photosensitivity and develop skin cancer at an approximately 2,000-fold increased frequency over non-XP patients. While all XP types share similar clinical phenotypes, particularly with respect to UV sensitivity and skin cancer, there is some disparity among groups. XPA patients experience some of the most severe symptoms of the disease. Neurodegeneration is also prevalent among XPA, XPB, XPD, and XPG patients. XPB and XPD patients are also at risk for the related disorders Cockayne syndrome (CS) and trichothiodystrophy (TTD)[63,64].

CS patients do not develop cancers at the extreme rates of XP patients. CS is primarily associated with premature aging, neurodegeneration, and UV sensitivity[64]. The disorder is primarily caused by mutations in *ERCC8* (encoding CSA protein) or *ERCC6* (encoding CSB protein). CSA and CSB are involved in TC NER.

A third NER-related disorder is TTD. While symptoms vary, the disease is characterized by brittle hair, dry skin, and developmental/neurological deficiencies[65]. About 75% of patients exhibit ichthyosis and about 50% exhibit photosensitivity[63,65]. TTD is caused by mutations in *XPB*, *XPD*, *TTDA*, or *TTDN1*. XPB, XPD, and TTDA are all subunits of TFIIH, involved in both transcription and damage verification during NER. The function of TTDN1 is less clear, and is associated with the non-photosensitive form of TTD.

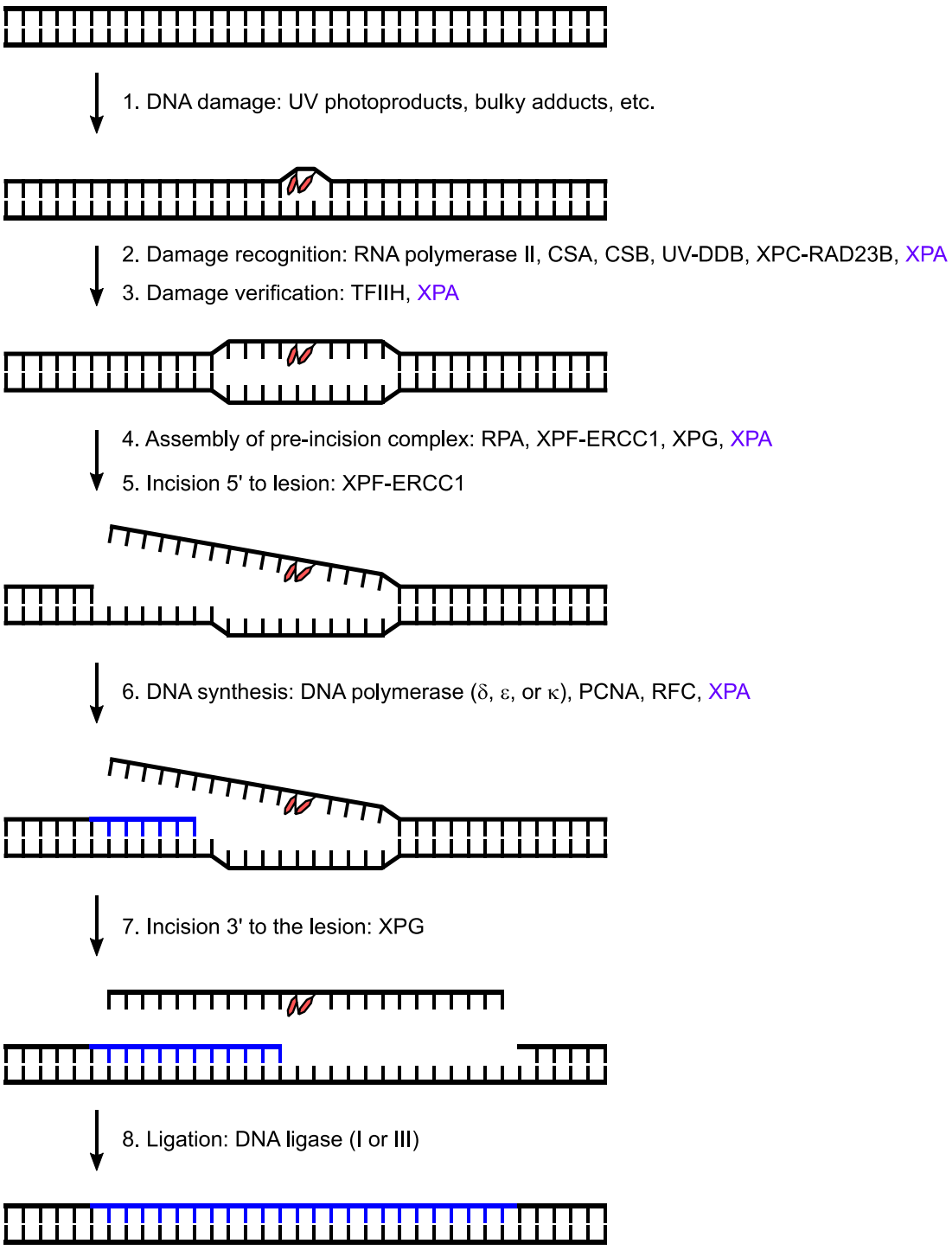

Figure 1. Summary of eukaryotic nucleotide excision repair.

Major changes to DNA during eukaryotic nucleotide excision repair are illustrated. The major proteins (not a comprehensive list) involved are indicated at the appropriate step. The potential involvement of XPA at various steps is indicated in purple.

1.3 XPA

1.3.1 Structure and Disorder

There is limited structural data for full-length XPA, due to large regions of conserved intrinsic disorder (Figure 2a), particularly in the N- and C-termini[66-69]. In the last 20 years, the study of intrinsically disordered proteins and protein domains has shown that lack of structure can actually be necessary for certain proteins to function[70-72]. Though diverse, many such disordered proteins are involved in molecular recognition[73]. Upon binding their target, disordered regions may fold and become structured[74] or remain flexible[75]. It has been suggested that the intrinsically disordered regions of XPA may play a role in binding to DNA and/or other NER proteins[69].

The human XPA protein contains 273 amino acids with a molecular weight of 31.4 kDa. The minimal DNA-binding domain (DBD) was first identified by Tanaka and colleagues[76], and later expanded to include residues 98-239 (ref. [77]), covering about half of the total protein length and including a zinc-finger motif[76,78-81]. Available structures are restricted to the DBD, including early solution NMR studies of human XPA[50,82,83] (Figure 2b), a recent crystal structure of the extended human DBD[84] (Figure 2c), co-crystal structures of yeast Rad14 (XPA homolog) on damaged DNA[85,86] (Figure 2d), and a cryo-electron microscopy structure of XPA bound to DNA with TFIIH[87] (Figure 2e). These structures all indicate the presence of a basic cleft, or cluster of positively charged residues (Figure 2c), presumably involved in binding the negatively charged backbone of DNA[68,83,88].

The Rad14 minimal DBD structures suggest that XPA binds as a dimer flanking the site of damage and produces a 70° bend in the DNA[85,86]. XPA stoichiometry, both on and off of DNA,

remains controversial. The Rad14 structure supports previous studies which concluded that XPA binds DNA as a homodimer[89,90]. Other reports indicate that XPA binds DNA as a monomer[68,87].

Finally, a combination of structural studies and biochemistry (EMSA[48], gel filtration chromatography[66], and circular dichroism[91]) have suggested that XPA binding to damaged DNA induces conformational changes in the protein that are associated with different binding modes. Although DNA bending was not measured directly in these studies, the observed binding modes may reflect the formation of protein-DNA complexes containing bent DNA and stably bound XPA. Recent molecular dynamics simulations of docking between XPA residues 98-210 (PDB 1XPA) and bent 10bp dsDNA containing a CPD (PDB 1N4E) shows formation of increased secondary structure in XPA compared to the free protein simulations[92].

1.3.2 XPA-DNA Interactions

In support of a damage recognition role for XPA, there is substantial evidence for the protein's specificity to several definitive substrates for NER. Some of the earliest studies demonstrated XPA's notable affinity for UV-irradiated DNA[48,55,76,79,93,94], with reported specificities as high as 1,000-fold[ii] (when calculated to account for non-specific bases in damaged

[ii] Note that while it is nearly impossible, in part due to the effects of experimental method and ionic strength/buffer conditions on XPA-DNA binding[90,95,96], to compare reported specificities between studies, I have attempted to do so in the most consistent manner possible. When the information is available, fold-specificities are reported as the difference in binding affinities between substrates—as published by the original authors—multiplied by the number of non-specific bases in the damaged substrate. In some cases, this is how the authors presented the

substrate)[93] over non-damaged dsDNA. While some groups report no specificity of XPA for CPD lesions[60,94], others have observed higher affinity for a CPD (~90-fold specificity) than non-damaged DNA[97]. Furthermore, in all studies, XPA had a significantly higher affinity for 6-4PPs than for both CPDs and non-damaged DNA. In a side-by-side comparison of XPA and XPC, an established recognition protein in NER, the two proteins exhibited similar fold-specificity (~75-fold) for a 6-4PP; XPC had a higher affinity for both damaged and non-damaged DNA, resulting in an analogous K_D ratio[60]. These data suggest that XPA prefers binding to a lesion which is more distorting to the DNA helix (i.e. to a 6-4PP, which induces a 44° bend, versus the less destabilizing CPD, which does not readily form a kinked structure)[27].

XPA binding to a variety of base adducts has also been of interest. XPA binds preferentially to AAF-adducted dsDNA over non-damaged dsDNA[85,86,90,98,99]. It does not, however, display this specificity when the AAF adduct is placed within a mismatched DNA bubble[100]. Furthermore, XPA has demonstrated specificity for a C8-aminofluorene (C8-AF) adduct and an N^2-acetylnaphthyl (N^2-AAN) adduct, although not to the same degree as AAF[86]. While all three adducts destabilize the DNA helix and induce bending, the authors suggest that the preferential binding to dG-C8-AAF, compared to dG-C8-AF or dG-N^2-AAN, might be due to the flexibility of the helix and the energy required to form a sharp bend when in complex with XPA; the two rings of the dG-C8-AAF lesion are in plane and able to intercalate/stack with adjacent bases, thereby stabilizing the helical kink[85,86].

---

data in their original reports.[46,93,94] For fold differences as reported in original studies, please refer to Table 1 (Appendix B). Also please note that not all authors provide quantitative comparisons for binding affinity.

XPA has increased affinity (up to ~250-fold reported[101]) for 1,2-GG cisplatin-adducted DNA[76,79,94,96,101-103] over non-damaged DNA, but less than 100-fold specificity for 1,3-GTG cisplatin-adducted DNA[101] and no specificity for a dinuclear analogue[102]. The 1,2-GG cisplatin intrastrand crosslink induces a rigid 30-35° bend in the DNA helix[104,105]. The 1,3-GTG cisplatin intrastrand crosslink induces a similar bend angle, but confers different thermodynamic properties to the helix[101]; the dinuclear analogue induces helical flexibility but not directional bending[106]. In an effort to test XPA affinity for different "rigid bends," Zou and colleagues studied binding to DNA substrates with two-, three-, or four-carbon tethers connecting adjacent guanine bases[100]. These intrastrand crosslinks induce 30°, 11.7°, or 7.4° bends in the DNA helix, respectively[107]. Interestingly, no specificity was reported for XPA binding to any of these substrates[100]. These data suggest that a bend alone is not sufficient to enhance XPA binding; however, the presence of the carbon tether may impair the ability of XPA to test for DNA bending and form stable complexes with even sharper bends.

One study reported no specificity of XPA for psoralen-treated dsDNA[94]. Psoralen is able to intercalate DNA, and upon UV treatment, forms covalent monoadducts at pyrimidine bases as well as diadduct interstrand crosslinks. A combination of studies have shown that the psoralen monoadduct has little impact on the DNA helix curvature or flexibility, and that while the diadduct does cause unwinding of the DNA about the lesion, it has little impact on helical secondary structure[108,109]. Additionally, XPA has demonstrated at least 50-fold specificity for dsDNA with a mitomycin C interstrand crosslink[91,110]. This lesion has also been reported to not significantly impact the DNA helix, but may cause some local bending or distortion[111]. It is possible that these low levels of specificity are only apparent under certain experimental conditions. Nonetheless,

compared to lesions which are known to cause significant DNA bending, XPA consistently demonstrates higher affinity.

Other DNA modifications have been investigated to gain further insight into the structural and thermodynamic requirements for XPA specificity. In one study, XPA demonstrated no specificity for a C4' pivaloyl DNA backbone adduct, which disrupts hydrogen bonding but does not distort the helix[98]. Others have reported preferential binding of XPA to dsDNA with a short bubble of three or four mismatched bases[98,100,102], or a single-stranded loop of three nucleotides inserted into one strand of duplex DNA[96,102], both of which do cause helical distortion. Additionally, XPA shows some specificity for DNA bases replaced with 5-nitroindole or 3-nitropyrrole nucleoside analogs that maintain all properties of B-form DNA, except for Watson-Crick hydrogen bonding[98,102]. Compared to these minor distortions, XPA binds avidly to three- and four-way dsDNA junctions, engineered to mimic helical kinks[102,103]. Although there has been no direct comparison of these substrates with NER lesions like AAF or 6-4PPs, the consensus appears to be that bent structures that maintain some amount of flexibility are recognized with the highest affinity by XPA.

Additional studies revealed that XPA also binds preferentially to partially single stranded DNA and forked substrates that more closely resemble unwound NER intermediates. XPA has significantly higher affinity—up to 120-fold specificity, reported by one group[77]—for forked Y-shaped substrates (i.e. ss/dsDNA junctions) than for non-modified dsDNA or ssDNA[100,102,103,112]. When compared to ds/dsDNA junctions, however, XPA has even higher affinity[102,103]. Furthermore, while these forked substrates could be considered as NER intermediates, they also represent structures that readily adopt a kinked DNA helix, and allow bending into the major groove, which might also mimic lesion substrates[85]. It is possible that, regardless of the step in the

NER pathway, XPA is able to recognize bent DNA structures and this property facilitates its role in damage recognition as well as in later steps. The common characteristic between "good" XPA substrates appears to be the ability to form a stable, sharply kinked conformation.

For a more comprehensive summary of published biochemical studies on XPA affinity for different DNA substrates, please refer to Table 1 (Appendix B). For further discussion of the energetics of DNA bending and protein binding, please refer to Appendix D.

1.3.3 XPA-Protein Interactions

In addition to having binding specificity for damaged DNA on its own, the XPA has a well-established role as a scaffold protein during DNA repair. A summary of major interactions reported in the literature is outlined in Table 2 (Appendix C). These have largely been studied using yeast two-hybrid, pull-down assays, and other biochemical methods. XPA's potential roles in multiple steps of the NER pathway are indicated by its interactions with proteins involved in essentially every step of repair (Figure 2a).

In support for an early role of XPA during damage recognition, interactions with both UV-DDB (via DDB2)[52,113] and XPC[114,115] have been reported. XPA and XPC do not appear to interact together on DNA, suggesting that a hand-off may occur between these two proteins during repair[116]. However, Matsunaga and colleagues showed that XPA and UV-DDB bind damaged DNA together. Moreover, this interaction increases the affinity of both proteins for a CPD lesion[113]. These data were used by Mattaparthi and colleagues in a computational study to show that the interaction between DDB2 and XPA residues 185-226 is likely strong and transient, involving 7-9 salt bridges and 16-20 hydrogen bonds[117].

XPA also likely plays an important role during damage verification; there is strong evidence supporting an interaction between XPA and TFIIH on DNA. Biochemical work[48,118,119] and a recent cryo-electron microscopy[87] structure show that XPA interacts with multiple TFIIH subunits: XPB, XPD, p52, and p8/TTDA. Furthermore, the presence of XPA enhances TFIIH specificity for NER substrates by promoting translocation along non-damaged DNA[61,62] but not at bulky cisplatin lesions[43,44].

The single-stranded binding protein RPA is thought to stabilize unwound DNA in the pre-incision complex. Extensive biochemical and structural work has shown that XPA and RPA interact both on and off of DNA[51,55,57,96,120-124]. Specifically, XPA interacts with the subunits RPA32 and RPA70, but not RPA14. Both of these interactions appear to be mediated by the N-terminus of XPA (upstream of the DBD). Mutations in XPA residues K141 and K179 impair binding to RPA70 but not to damaged DNA[56]. Furthermore, several groups report that RPA enhances the affinity of XPA binding to short damaged dsDNA substrates[55,97,102], although there is some disagreement about the cooperativity of binding[90,96].

XPA interaction with the ERCC1-XPF endonuclease complex also places XPA in the pre-incision or incision complex[125]. XPA is required for the recruitment of ERCC1/XPF to damaged DNA, and this appears to be mediated via direct interaction with residues 91-119 of ERCC1[49]. In a filter binding assay, XPA demonstrated enhanced affinity for UV-irradiated DNA, but not non-damaged DNA, in the presence of purified ERCC1[126]. A ternary complex of ERCC1 with XPA and RPA has also been observed[122]. Several other XPA-protein interactions have been reported with less clear implications for NER. These are included in Appendix C.

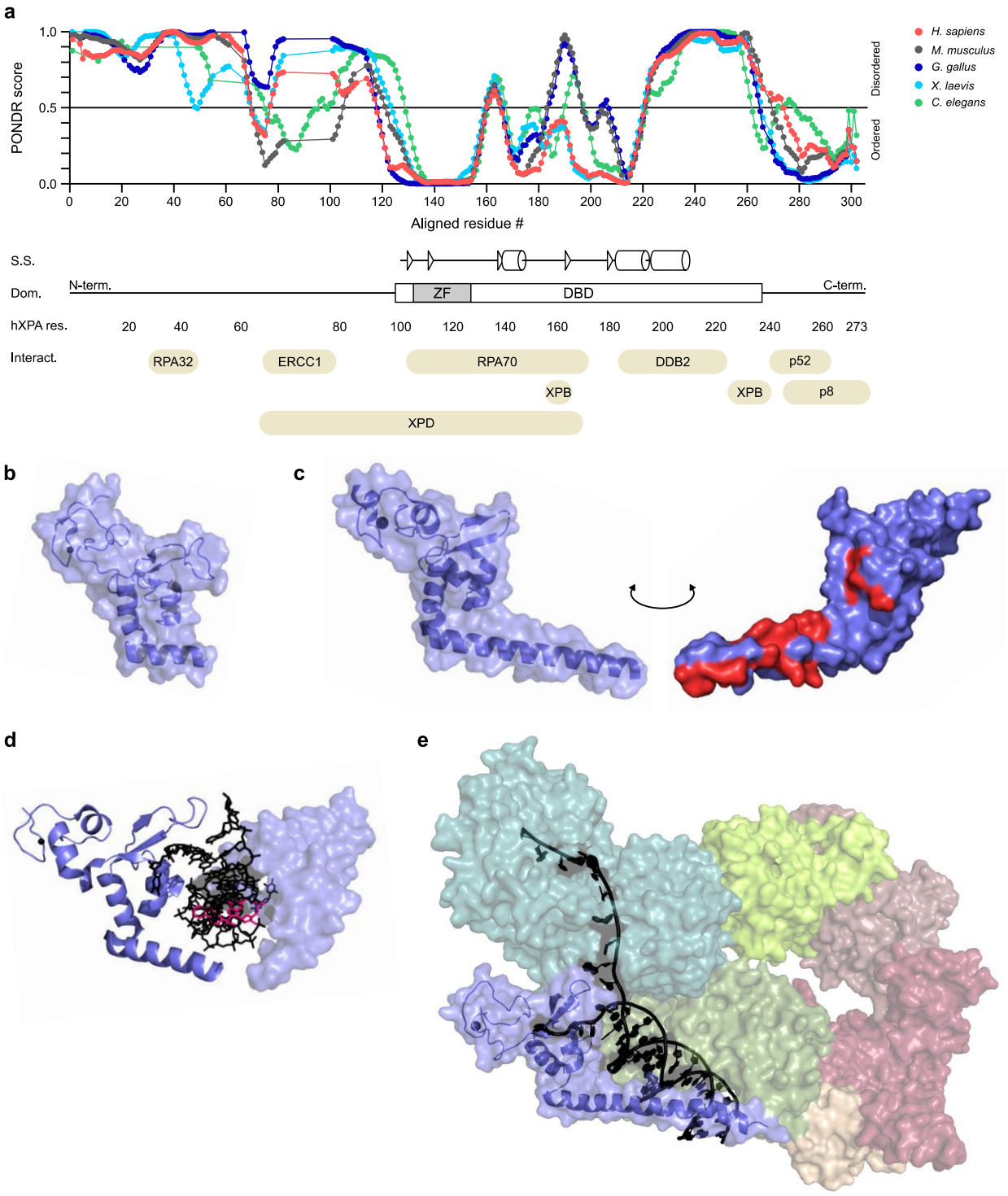

Figure 2. XPA structure and disorder.

[a] Protein sequences of XPA homologs from multiples species were aligned using PROMALS3D[127] and disorder predictions for each sequence were obtained via the PONDR VL-XT algorithm[128]. S.S., secondary

structure elements (triangles, beta sheets; cylinders, alpha helices) based on the NMR structure of the human XPA DBD (PDB 1XPA). Dom., conserved domains of human XPA protein (ZF, zinc finger; DBD, DNA-binding domain). Interact., some published interactions between XPA and NER proteins (see Appendix C); tan ovals represent XPA residues involved. [b-e] Major resolved structures of XPA and Rad14 DNA-binding domains (blue). [b] Solution NMR structure of the minimal DNA binding domain of human XPA (M98-N210) without DNA. PDB 1XPA. [c] Crystal structure of extended DNA binding domain of human XPA (M98-R231) without DNA. Right, structure is rotated and positively charged residues of basic cleft are colored red. PDB 6J44. [d] Co-crystal structure of yeast Rad14 DNA binding domain bound to DNA containing an AAF adduct (magenta). PDB 5A3D. [e] Cryo-electron microscopy structure of XPA bound (full-length protein used, residues I104-R237 resolved) to DNA with TFIIH. TFIIH subunits: p8 (wheat), XPB (sage), XPD (teal), p44 (lime), p34 (mauve), p52 (raspberry). PDB 6RO4.

1.4 Hypotheses and Scope

Motivated by compelling reports in the literature that XPA is able to recognize DNA damage, and further encouraged by results obtained during this dissertation, we set out to resolve five fundamental issues regarding how XPA interacts with DNA and its mechanism of damage search.

1. Does XPA have specificity for NER substrates? Numerous bulk studies have shown that XPA binds preferentially to DNA containing an AAF adduct or UV-photoproduct. We sought to demonstrate this specificity at the single molecule level. Atomic force microscopy (AFM) was used to show that XPA binds more frequently at an AAF adduct and single molecule fluorescence microscopy demonstrated that XPA pauses at UV-lesions in long DNA molecules.

2. What is the stoichiometry of XPA binding to DNA? Based on the Rad14 co-crystal structure and several biochemical reports, we predicted that XPA would bind DNA lesions as a

homodimer. However, we also noted some potential flaws in the previous studies, including the use of short DNA substrates and artificially high protein concentrations. To answer the question of stoichiometry as directly as possible, we turned to atomic force microscopy, as it is uniquely able to distinguish between true dimer complexes and distinct binding events on the same DNA molecule[129]. We found that XPA binds DNA as a monomer at both damaged and non-damaged sites.

3. What role does DNA bending have in damage search? Again, based on the Rad14 co-crystal structure and the well-known importance of DNA bending for other repair proteins, we hypothesized that XPA would induce DNA bending as part of its target search. Atomic force microscopy offers a direct measure of DNA bend angles at specific sites (i.e. at a lesion or bound protein). As predicted, we found that (1) XPA has specificity for DNA lesions (in this case, AAF) that induce DNA bending and (2) that XPA bends DNA even further at both non-damaged and damaged sites.

4. What modes of diffusion does XPA exhibit? We then sought to identify search strategies used by XPA, as discussed in Section 1.1. In order to gain insight into the dynamics of XPA damage search, we turned to single molecule fluorescence microscopy via the DNA tightrope assay. This was used to distinguish between three-dimensional and one-dimensional diffusion as well as well different modes of one-dimensional diffusion (short-range sliding and long-range hopping). We tested dose-dependent effects of DNA damage on XPA diffusive behavior.

5. What role do the disordered N- and C-termini play in XPA-DNA interactions? Finally, we hypothesized that each mode we observed during XPA damage search on DNA corresponded to a distinct conformational state. Furthermore, we predicted that the intrinsically disordered N- and C-terminal arms of XPA participated in damage recognition and changing between

conformational states. To test this, DNA tightrope experiments were performed with a truncated form of XPA.

Finally, we propose a model of XPA episodic motion in which different conformational states of the protein are associated with different modes of DNA target search and the presence of helix-distorting DNA damage stabilizes tighter binding.

1.5 Approach: Single Molecule Methods for Studying Protein-DNA Interactions

In order to address these questions and elucidate how XPA recognizes DNA damage we used two single molecule techniques. First, atomic force microscopy (AFM) was used to assess binding specificity, stoichiometry, and DNA bending. To accomplish this, a new PeakForce Tapping® AFM mode (Bruker) was validated for calculating the molecular weight of small proteins bound to DNA. Second, we used single molecule fluorescence microscopy to follow how quantum dot-labeled XPA interrogates DNA for damage in real time. Single molecule approaches offer unique advantages over bulk studies, discussed below and throughout this dissertation. A brief introduction to these methods is provided below. For further discussion on the uses, advantages, and limitations of AFM and the DNA tightrope assay, please refer to Appendix D (Studying Protein-DNA Interactions Using Atomic Force Microscopy), Appendix E (Dancing on DNA tightropes: Watching Repair Proteins Interrogate DNA in Real Time), and Appendix F (Single-Molecule Methods for Nucleotide Excision Repair: Building a System to Watch Repair in Real Time).

1.5.1 Atomic Force Microscopy[iii]

Developed in the mid-1980's, the atomic force microscope has become an increasingly powerful instrument for studying physical properties of materials on an atomic scale[130]. When studying protein-DNA interactions, it is important to keep in mind the forces that govern them. The four major interactions are: (1) hydrogen bonding between side chain and main chain amino acids and the floor of the major or minor groove of the DNA helix, (2) ionic interactions between the negatively charged phosphate backbone of DNA and positive amino acid side chains, (3) hydrophobic interactions and particularly pi stacking of DNA bases and aromatic side chains, and (4) Van der Waals forces over large surface areas[131,132].

The first reports of atomic force microscopy (AFM) imaging of protein-DNA complexes were in 1992 of the *E. coli* RNA polymerase complexed with DNA[133] and of DNA polymerase on M13 phage DNA[134]. Since then, AFM has proven to have unique advantages in the study of protein-nucleic acid interactions. AFM imaging is relatively simple and the process allows for samples to remain under more physiological conditions. Relatively long DNA substrates may be used and there is no requirement for labeling, staining, or fixation of either the DNA or the protein. Most importantly, it provides direct imaging at the single molecule level, and thus rare events can be observed that would otherwise by obscured in bulk biochemistry techniques.

[iii] This section (1.5.1) has been adapted from ref. [129]. Please find the full text in Appendix D. Also refer to Appendix F for further discussion of AFM of nucleoprotein complexes.

1.5.2 DNA Tightrope Assay[iv]

In order to understand how DNA repair proteins find damaged sites in a vast excess of non-damaged DNA, the field of DNA repair has moved to various single molecule approaches allowing direct visualization of proteins interacting with their DNA substrates[136]. These single molecule techniques can provide unique insights into population trends without losing detailed information on individual particles or events[137]. An optical platform consisting of DNA tightropes was developed by Neil Kad at the University of Vermont and first used to study bacterial nucleotide excision repair (NER) proteins[138,139] and base excision repair (BER) glycosylases[140]. This DNA tightrope assay takes a similar approach to the DNA curtain setup developed by Dr. Eric Greene and colleagues[18,141] with one important difference. The tightrope itself is established by suspending long molecules of dsDNA (~90% contour length) between poly-L-lysine coated 5 micron beads dispersed in a flow cell. Visualizing repair proteins of interest up off the surface requires labels with bright fluorescent signals, and real-time imaging requires photostability over long periods. To accomplish these two needs, repair proteins are conjugated to quantum dots (Qdots) with appropriate antibodies and added to the flow cell. Interactions are recorded in real time, in the absence of flow, using oblique angle fluorescence on a total internal reflection fluorescence (TIRF) microscope with a CMOS or EECD camera[142].

[iv] This section (1.5.2) has been adapted from ref. [135]. Please find full text in Appendix E. Also refer to Appendix F for further discussion of DNA tightrope assay.

2.0 Materials and Methods

2.1 Protein Purification

2.1.1 His-flXPA

Full length, wild-type human XPA cDNA was cloned into the pIBA35 vector with an N-terminal His tag. The plasmid was transformed into One Shot BL21(DE3)pLysS competent *E. coli* cells (Invitrogen). Cultures were grown in LB medium containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol at 37°C until the OD_{600} reached ~0.6. At this point, expression was induced with 0.5 mM IPTG and 10 μM $ZnCl_2$ and cultures continued to grow for 4 hours. Cells were harvested by centrifugation at 16,000 × *g* for 8 minutes at 4°C. All the following purification steps were performed on ice or at 4°C. Cell pellets were resuspended in His-XPA lysis buffer (25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM DTT, 10% (v/v) glycerol, 10 μM $ZnCl_2$, and 30 mM imidazole), lysed by sonication, and the insoluble fraction was pelleted by centrifugation at 45,000 × *g* for 45 minutes. The supernatant was loaded onto an equilibrated HisTrap HP nickel column (GE) and washed with 30 column volumes (CV) of His buffer A (25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM DTT, 10% glycerol, 10 μM $ZnCl_2$, and 30 mM imidazole). The sample was eluted with a gradient of 0-100% in 10 CV, His buffer A to His buffer B (25 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM DTT, 10% glycerol, 10 μM $ZnCl_2$, and 500 mM imidazole). Fractions containing XPA were pooled and loaded onto an equilibrated MonoQ column and washed with 5 CV MonoQ buffer A (25 mM Tris-HCl, pH 7.5, 50 mM NaCl, 5 mM DTT, 10% glycerol, and 10 μM $ZnCl_2$). XPA was eluted with a gradient of 0-100% in 15 CV MonoQ buffer A to MonoQ

buffer B (25 mM Tris-HCl, pH 7.5, 1 M NaCl, 5 mM DTT, 10% glycerol, and 10 μM $ZnCl_2$). Fractions containing XPA were pooled, diluted with MonoQ buffer A, loaded onto an equilibrated Heparin column, and washed with 5 CV MonoQ buffer A. XPA was eluted with a gradient of 0-100% in 35 CV MonoQ buffer A to MonoQ buffer B. Fractions containing XPA were pooled and loaded onto a size exclusion column (HiLoad 16/60 Superdex 200), which was equilibrated with the His-XPA SEC buffer (50 mM HEPES, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 5% glycerol, and 10 μM $ZnCl_2$). XPA eluted as a single peak and peak fractions were pooled, aliquoted, and flash frozen in liquid nitrogen and stored at -80°C.

2.1.2 His-flXPA-StrepII and His-truncXPA-StrepII

Full length, wild-type human XPA cDNA was cloned into the pIBA43 vector with an N-terminal His tag and C-terminal StrepII tag. To make truncated XPA mutant, the N- and C-termini were deleted from this plasmid (cloning by Gene Universal), leaving only residues 98-239. Both constructs were expressed and purified the same way.

The plasmid was transformed into One Shot BL21(DE3)pLysS competent *E. coli* cells (Invitrogen). Cultures were grown in LB medium containing 100 μg/ml ampicillin, 34 μg/ml chloramphenicol, and 10 μM $ZnCl_2$ at 37°C until the OD_{600} reached ~0.4. The temperature was then decreased to 16°C and growth was continued until an OD_{600} of ~0.6 was achieved. At this point, expression was induced with 0.2 mM IPTG and cultures continued to grow overnight. Cells were harvested by centrifugation at 5,000 × *g* for 30 minutes at 4°C. All the following purification steps were performed on ice or at 4°C and, between each step, samples were analyzed on 4-12% Bis-Tris SDS gels and stained with SimplyBlue SafeStain (Invitrogen). Cell pellets were resuspended in His-XPA-StrepII lysis buffer (50 mM potassium phosphate, pH 7.9, 200 mM KCl,

20 mM imidazole, 5% glycerol, 0.02% sodium azide, and EDTA-free Protease inhibitor cocktail), lysed by sonication, and the insoluble fraction was pelleted by ultracentrifugation at 147,000 × g for 2 hours. The supernatant was loaded onto an equilibrated HisTrap HP nickel column (GE) and washed with 30 CV His buffer C (3.5 mM KH_2PO_4, 46.5 mM K_2HPO_4, 200 mM KCl, 20 mM imidazole, 5% glycerol, and 0.02% sodium azide). The sample was eluted with a gradient of 0-50%, His buffer C to His buffer D (3.5 mM KH_2PO_4, 46.5 mM K_2HPO_4, 200 mM KCl, 500 mM imidazole, 5% glycerol, and 0.02% sodium azide). Fractions containing XPA were pooled and loaded onto an equilibrated StrepTrap HP column with StrepTactin sepharose (GE) and washed with 50 column volumes Strep buffer A (50 mM Tris-HCl, pH 8.0, 200 mM KCl, and 0.02% sodium azide) and eluted with Strep buffer B (50 mM Tris-HCl, pH 8.0, 200 mM KCl, 5 mM desthiobiotin, and 0.02% sodium azide). Fractions containing XPA were pooled and dialyzed into His-XPA-StrepII SEC buffer (50 mM HEPES, pH 8.0, 200 mM KCl, 5 mM $MgCl_2$, 5 mM DTT, 10% glycerol, and 0.02% sodium azide). Size exclusion chromatography was performed using the AKTA FPLC on a HiLoad 16/60 Superdex 200 column (Amersham). XPA eluted as a single peak and peak fractions were pooled, aliquoted, and flash frozen in liquid nitrogen and stored at -80°C. His-XPA-StrepII and His-XPA were compared by EMSA and AFM to confirm that the addition of the StrepII tag did not affect protein behavior.

His-flXPA was used in all AFM experiments and some DNA tightrope experiments. His-flXPA-StrepII was used for all electrophoretic mobility shift assays, multiangle light scattering, and some DNA tightrope experiments. To verify that both protein preparations exhibited similar behavior, they were compared by AFM with respect to the following parameters: binding position on AAF_{538}, induced DNA bend angle, and AFM volume (Figure 3a-d). Because His-flXPA-StrepII has a slightly higher molecular weight than His-flXPA, His-flXPA-StrepII results were only used

for validation and were not combined with His-flXPA results for AFM experiments. Since both preparations behaved similarly in our AFM studies, we combined the data obtained in the DNA tightrope experiments.

2.2 Multiangle Light Scattering

Multiangle light scattering combined with size exclusion chromatography (SEC-MALS) of purified His-flXPA-StrepII in His-XPA-StrepII SEC buffer was performed as previously described[143,144].

2.3 DNA Substrate Preparation

2.3.1 AAF_{37} Oligo

The 37 nt oligonucleotide containing one dG-C8-AAF lesion (AAF_{37}-top, sequence below) was a gift from Thomas Carell, prepared as published[85].

2.3.2 37 bp DNA Duplexes for EMSA

ND_{37} was prepared by annealing ND_{37}-top and FAM_{37}-bottom. AAF_{37} was prepared by annealing AAF_{37}-top and FAM_{37}-bottom. CPD_{37} was prepared by annealing CPD_{37}-top and CPD_{37}-bottom. Annealing reactions contained 1.25 μM top strand, 1 μM bottom strand, 10 mM Tris-HCl, pH 8.0, and 100 mM KCl. Reactions were incubated at 95°C for 5 minutes then cooled

slowly to room temperature Oligonucleotide sequences (all purchased from IDT except AAF_{37}-top; G^{AAF} = dG-C8-AAF; T<>T = CPD; 6FAM = fluorescein):

ND_{37}-top: 5'-phosphate-CCGAGTCATTCCTGCAGCGAGTCCATGGGAGTCAAAT

AAF_{37}-top: 5'-phosphate-CCGAGTCATTCCT**G**$^{\mathbf{AAF}}$CAGCGAGTCCATGGGAGTCAAAT

CPD_{37}-top: 5'-phosphate-CCGAGTCATTCCTGCAGCGA**T<>T**CCATGGGAGTCAAAT

FAM_{37}-bottom: 5'-6FAM-ATTTGACTCCCATGGACTCGCTGCAGGAATGACTCGG

CPD_{37}-bottom: 5'-6FAM- ATTTGACTCCCATGGAATCGCTGCAGGAATGACTCGG

2.3.3 Defined Lesion Plasmids

Plasmids containing single site-specific dG-C8-AAF adducts were prepared as described previously (see Appendix F for a detailed protocol)[21,142]. Briefly, purified pSCW01 plasmids were nicked by Nt.BstNBI to create a 37-base gap. A 37mer containing a single dG-C8-AAF (AAF_{37}-top, above) was annealed into this gap and the backbone was sealed with T4 DNA ligase.

2.3.4 DNA Duplexes for AFM

Substrates for AFM were prepared as described previously (see Appendix F)[142]. Essentially, a 538 bp DNA fragment was cut out of either unmodified pSCW01 plasmid (for ND_{538}) or pSCW01 with a site-specific dG-C8-AAF lesion, described above (for AAF_{538}). The plasmid was incubated with restriction enzymes XmnI and PciI, cutting 372 bp 5' to and 165 bp 3' to the lesion, respectively. $Nick_{514}$ was prepared by amplifying a 514 bp fragment from the pSCW01 plasmid and treating with Nt.BspQI to create a nick at 36% of the DNA contour length.

2.3.5 Long DNA Substrates for Tightrope Assay

The NDλ substrate was prepared by diluting λ genomic DNA (NEB) to 50 ng/µl in 10 mM Tris-HCl, pH 8.5. The $UV\lambda_{20J}$ substrate was prepared by treating NDλ with 20 J/m^2 of UV-C radiation (254 nm). A qPCR assay was performed previously to confirm the presence of UV photoproducts (6-4 PPs and CPDs) at a density of ~1 lesion per 2.2 kbp[21]. To prepare $UV\lambda_{80J}$ and increase the lesion density such that there was ~1 lesion per 550 bp (i.e. 1 6-4PP every 2.2 kbp), NDλ was treated with 80 J/m^2 of UV-C. The dG-C8-AAF arrays were prepared as described previously (see Appendix F)[21,142], using the defined lesion plasmid described above. Lesion-containing pSCW01 was linearized via restriction digest by XhoI (NEB) then incubated with T4 DNA ligase (NEB) to achieve long (> 40 kbp) tandemly ligated products with one dG-C8-AAF every 2 kbp.

2.4 Electrophoretic Mobility Shift Assay

XPA-DNA reactions were prepared by combining 8 nM 37 bp DNA with varying amounts of His-flXPA-StrepII in XPA EMSA buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM DTT, 0.5 mg/ml BSA, and 5% glycerol) in a final reaction volume of 10 µl. Each reaction was incubated for 25 minutes at room temperature then immediately loaded on two pre-run 5% non-denaturing polyacrylamide gels (37.5:1 acrylamide:bis). Both pre-run and run were performed at 4°C, in 0.5X TBE buffer (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA, pH 8.4), at constant voltage (90 V). DNA bands were visualized using a laser scanner for fluorescence (Typhoon, Amersham).

Gel images were quantified by measuring signal intensities of each band (ImageJ, NIH). The percentage of DNA bound was determined by dividing the intensity of the shifted ("bound") DNA by the sum of all bands in a lane. These values were plotted against XPA concentration and the data were fit to the following equation via nonlinear regression (GraphPad Prism):

$$\%\ DNA\ Bound = 100\ \times \frac{(P + D + K_D) - \sqrt{(P + D + K_D)^2 - 4PD}}{2D} \quad \textbf{Equation 1}$$

where K_D is the equilibrium dissociation constant, P is the total protein concentration, and D is the total DNA concentration. This model was chosen because our experimental conditions required that the DNA concentration be in the same molar range as the K_D (ref. [145,146]).

2.5 Atomic Force Microscopy

2.5.1 Sample Preparation

Samples for AFM were prepared as previously described (see Appendix F)[142]. All buffers and solutions were first filtered through 0.02 μm sterile filters (Whatman). For imaging of free proteins or free DNA (i.e. no reaction), the sample was diluted to either 40 nM (protein) or 4 nM (DNA) in AFM deposition buffer (25 mM HEPES, pH 7.5, 25 mM NaOAc, and 10 mM $Mg(OAc)_2$) which had been pre-warmed to 65°C and brought back to room temperature. XPA-DNA reactions consisted of 100 nM 538 bp DNA (ND_{538} or AAF_{538}) and 0.6-4 μM His-flXPA in XPA AFM buffer (50 mM HEPES, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 10 μM $ZnCl_2$, 5 mM DTT, and 5% glycerol) in a total volume of 10 μl. APE1-DNA and Polβ-DNA reactions consisted of 100 nM 514 bp DNA ($Nick_{514}$) and 500 nM protein in APE1 buffer (50 mM HEPES, pH 7.5,

150 mM NaCl, 10 mM $MgCl_2$) in a total volume of 10 µl. Each binding reaction was incubated for 30 minutes at room temperature then diluted 1:25 in AFM deposition buffer. 25 µl droplets were deposited on freshly cleaved mica, allowed to equilibrate for 30 seconds with gentle rocking, then washed with 1 ml of filtered H_2O and dried under a gentle stream of nitrogen gas.

2.5.2 Data Collection

All AFM images were obtained using ScanAsyst PeakForce Tapping mode in air on a Multimode V Microscope with an E scanner (Bruker). Samples were scanned with a triangular tip with a nominal radius of 2 nm, mounted on a silicon nitride cantilever (SCANASYST-AIR, Bruker). Probes were replaced for each new experiment or more frequently as needed. 1 × 1 micron images were collected at a resolution of 512 × 512 pixels and a scan rate of 0.977 Hz. Peak force setpoint was 0.01988 V.

2.5.3 Data Analysis

2.5.3.1 Free protein standard

To generate the standard curve relating AFM volumes to molecular weight, analysis was performed on AFM images with the isolated protein samples. The following proteins of known MW were used: recombinant human HMGB1 (Abcam), His-tagged human APE1 (gift from Sam Wilson), His-tagged human DNA polymerase β (gift from Sam Wilson), and His-tagged UvrD (purified as published[147]). Particle dimensions were measured using Image SXM software and used to calculate volumes:

$$V = A \times (H - B)$$ **Equation 2**

where V is the particle volume, A is the area of the particle footprint (determined via a set density threshold above background noise), H is the mean height of the particle, and B is the background height of the overall image[148].

2.5.3.2 Intrinsic DNA bend angle

To determine the intrinsic bend angles of DNA substrates at 30% from each end, AFM images containing only the DNA were analyzed. DNA molecules used in this analysis had to be completely visible and isolated (i.e. not continuing past the edge of the image nor overlapping with itself or another molecule) and the total contour length must be within the range of ± 10% of the expected length. Measurements were done on TIF images using ImageJ software (NIH). The total DNA contour length was first measured and points at 30% from both ends were marked. Local bend angles at these sites were measured and are reported as the supplementary angle, θ (Figure 11).

2.5.3.3 Protein-DNA complexes

In addition to the criteria for usable DNA molecules (above) analysis of protein-DNA complexes first required the identification of bound proteins using the following criteria: (a) the height of the complex must be greater than the average height of the DNA molecule and (b) the complex width must be greater than the average width of the DNA molecule.

Methods for measuring protein binding position and induced DNA bend angle using ImageJ software (NIH) have been described in detail[142]. Briefly, binding position was determined by dividing the contour length of the DNA molecule from the center of a bound protein to the closest DNA end by the total DNA contour length. XPA-induced DNA bend angles were measured

at sites of bound XPA. In cases where two or more proteins were bound to the same DNA molecule, angles were not measured.

Specificity calculations from protein binding positions were performed as published by Erie and colleagues[149]. Histogram showing distribution of protein binding position between 0 and 50% contour length of AAF_{538} was fit by least squares nonlinear regression: $y = 3.6 + 10.4e^{-0.5\left(\frac{x-34.3}{9}\right)^2}$. Specificity (*S*) was calculated as:

$$S = N * \frac{A_{specific}}{A_{non-specific}} + 1 \qquad \textbf{Equation 3}$$

where *N* is the number of potential binding sites using an estimate of 8 bp (ref. [150]) for the DNA footprint of XPA (*N* = 538 bp – 8 bp + 1 = 534 bp) and $A_{specific}$ and $A_{non\text{-}specific}$ are the areas under the curve representing specific and non-specific binding, respectively (Figure 6).

To measure protein volume when bound to DNA, the DNA volume was estimated and subtracted from the total complex volume (Figure 9a). Image SXM software was used to trace the perimeter of the complex. The length of the DNA through this space was projected assuming that the DNA runs through the center of the complex. Then, two unbound regions of DNA on either side of the complex, with lengths corresponding to that of the complex, were delineated. In cases where the protein was bound near the end of the DNA or near another protein, two unbound regions of DNA were chosen at other available locations on the same molecule. Volumes of all three regions (complex, DNA1, and DNA2) were determined as above (Equation 2). Protein volume was determined as the total complex volume minus the average of the two unbound DNA volumes:

$$V_{protein} = V_{complex} - \frac{V_{DNA1} + V_{DNA2}}{2} \qquad \textbf{Equation 4}$$

Histograms of all AFM results were plotted and Gaussians were fit to the data by nonlinear regression in GraphPad Prism. The number of histogram bins correspond to the square root of the sample size.

2.6 DNA Tightrope Assay

2.6.1 Flow Cell Set-Up

All steps for reagent/material preparation, flow cell set-up, protein labeling, imaging, and data analysis for the tightrope assay have been described in detail (Appendix F)[142] according to methods developed previously[138,140]. Briefly, flow cells were prepared by attaching slides with inlet/outlet tubing to PEGylated coverslips via tape spacers. Flow cells were incubated in blocking buffer (10 mM HEPES, pH 7.5, 50 mM NaCl, and 1 mg/ml BSA) for 10 minutes, then poly-L-lysine coated silica microspheres (5 μm diameter) were flowed in and dispersed across the coverslip. Long DNA substrates were suspended between beads using continuous hydrodynamic flow with alternating direction in tightrope buffer (20 mM HEPES, pH 7.5, 50 mM NaCl, 3 mM $MgCl_2$).

2.6.2 Protein Labeling

His-flXPA, His-flXPA-StrepII, or His-truncXPA-StrepII was labeled with either 705 nm or 605 nm quantum dots (Qdots). The former strategy was accomplished by first incubating streptavidin-coated 705 Qdot (Invitrogen) with biotinylated anti-His antibody, at a final

concentration of 167 nM Qdot and 833 nM antibody. Then, this mixture was incubated with an equal volume of 167 nM XPA; the final XPA concentration was 83.3 nM. The latter strategy was accomplished by first incubating XPA with a mouse monoclonal anti-His antibody, both at a final concentration of 200 nM. Then, this mixture was incubated with an equal volume of 1 μM 605 Qdot conjugated to an anti-mouse secondary antibody (Invitrogen); final XPA concentration was 100 nM. Labeled protein mixtures were diluted 4:100 in XPA tightrope buffer (25 mM HEPES, pH 8.3, 100 mM KCl, 1 mM EDTA, 0.545 mg/ml BSA, 1 mM DTT, and 10% glycerol). The final XPA concentration in the flow cell was 3-4 nM. To check for background binding by the Qdots or antibodies, controls were performed using the above conjugations with buffer instead of XPA. To verify that the two different Qdot labeling strategies did not impact results, the behavior of XPA labeled with either the 605 Qdot or 705 Qdot was compared on UVλ_{20J} tightropes, showing no significant difference (Figure 3e).

For 150 mM NaCl experiments, 150 mM NaCl buffer (25 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM DTT, and 0.1 mg/ml BSA) was used in place of XPA tightrope buffer. For 1 M NaCl experiments, stationary particles of XPA were recorded in XPA tightrope buffer and, during recording, the buffer in the flow cell was replaced with 1 M NaCl buffer (1 M NaCl, 20 mM HEPES, pH 8.3, 80 mM KCl, 0.8 mM EDTA, and 8% glycerol), taking ~40 s of flow.

2.6.3 Data Collection

Movies of XPA-DNA interactions were recorded on an inverted fluorescence microscope (Nikon Ti) with 100X oil-based high-NA objective for TIRF-M and high-speed sCMOS camera (Andor). Qdots were excited with a 488 nm laser at an optimal oblique (sub-TIRF) angle and

visualized without an emission filter. Movies were taken for 5 minutes with frame rates between ~10 and ~12.5 fps.

2.6.4 Data Analysis

Movie files were converted to a time series of individual TIF files (NIS-Elements, Nikon) and imported into ImageJ (NIH). Kymographs were generated using the slice function over the trajectory of the particle along the DNA. These were processed by FFT bandpass filtering to reduce noise (filter range 3-40 pixels, with suppression of vertical stripes). A Gaussian Fit plugin was used to fit the fluorescence intensity in the kymograph to a one-dimensional Gaussian at each point along the x-axis (i.e. each frame or time point)[138]. Fitting data was processed using custom scripts in MATLAB (MathWorks) to exclude poorly fitted positions and convert particle position from pixels to nm.

First, each particle (i.e. one kymograph, 5 minute observation) was categorized based on whether it moved at all during the observation window (stationary vs. motile) and whether it dissociated during recording (persistent vs. dissociated). Dissociation was defined as the disappearance of Qdot-XPA for at least 200 frames (~20 seconds). We have previously reported that the mean positional accuracy for a 605 nm Qdot bound to biotin is 6 ± 3 nm by Gaussian fitting of the fluorescence intensity to a point spread function[21]. The position uncertainty over time has been determined to be 36 ± 3 nm (~100 bp), accounting for stage drift, DNA movement, and thermal fluctuations[21]. We used a conservative cutoff of 130 nm (three pixels, ~400 bp) to classify motile particles[22,23].

Motile particles were analyzed further for different modes of diffusion. Each kymograph was broken down into phases, falling into three possible modes of behavior: paused (particle

displacement not varying more than 130 nm), short-range diffusion (displacement between 130 nm and 690 nm), and long-range diffusion (displacement greater than 690 nm). Shorter range modes were only allowed to interrupt longer range modes if they persisted for at least 5 seconds. For example, if a particle was exhibiting short-range behavior, paused (i.e. stationary) for 10 seconds, then went back to short-range behavior, this would be counted as three phases; if the pause only lasted 2 seconds, this would be counted as a single short-range phase. If a particle was paused prior to recording, the first phase of a kymograph may be less than 5 seconds. Each motile particle was analyzed with respect to the following parameters: position range, phase switch rate, lifetime of each phase, and number of pause sites.

The mean squared displacement (MSD) was calculated for all motile phases (short-range and long-range) using custom scripts in MATLAB:

$$MSD(n\Delta t) = \frac{1}{N-n}\sum_{i=1}^{N-n}(x_{i+n} - x_i)^2 \qquad \textbf{Equation 5}$$

where N is total number of frames in the phase, n is the number of frames at a given time step, Δt is the time increment of one frame, and x_i is the particle position in the ith frame[151]. The diffusion coefficient (D) was determined by fitting a linear model of one-dimensional diffusion to the MSD plots:

$$MSD(n\Delta t) = 2D(n\Delta t) + y \qquad \textbf{Equation 6}$$

where y is a constant (y-intercept). Fittings resulting in R^2 less than 0.8 or using less than 10% of the MSD plot were not considered.

2.6.5 Calculation of Theoretical Constants

Calculations to determine the theoretical limit of the diffusion coefficient and the energy barriers to free diffusion were pursued as described[22,138]. All calculations were done based on the streptavidin-coated 705 Qdot labeling strategy, although similar results are obtained for 605 Qdot conjugated to a secondary antibody. First, the hydrodynamic radii of full-length human XPA (3.3 nm, ref. [94]) and the streptavidin-coated 705 Qdot (12.8 nm, ref. [152]) were used to estimate the hydrodynamic radius of Qdot-labeled XPA (R_{eff} = 12.873 nm).

Treating the labeled protein as a sphere allows us to define the diffusion coefficient (D) with the Stokes-Einstein equation:

$$D = \frac{k_B T}{\xi} \qquad \textbf{Equation 7}$$

where k_B is the Boltzmann constant (1.38 J/K), T is the temperature (298 K), and ξ is a friction term. The friction term for a protein sliding along DNA following the corkscrew path of the helix, described by Schurr[153] and modified slightly[154], is defined as:

$$\xi = 6\pi\eta R_{eff} + \left(\frac{2\pi}{10.5 \times BP}\right)^2 \left(8\pi\eta {R_{eff}}^3 + 6\pi\eta R_{OC} {R_{eff}}^2\right) \qquad \textbf{Equation 8}$$

where η is the viscosity of the medium (0.89×10^{-2} poise), BP is the distance between two DNA base pairs (0.34 nm), and R_{OC} is the off-center distance from the protein center of mass to the DNA helical axis (R_{eff} + 1 nm = 13.873 nm). Combining Equation 7 and Equation 8 permits calculation of the diffusion coefficient of Qdot-labeled XPA sliding along DNA with no energy barrier, or the theoretical limit to the diffusion coefficient (D_{lim}):

$$D_{lim} = \frac{k_B T}{6\pi\eta R_{eff} + \left(\frac{2\pi}{10.5 \times BP}\right)^2 \left(8\pi\eta {R_{eff}}^3 + 6\pi\eta R_{OC} {R_{eff}}^2\right)} \qquad \textbf{Equation 9}$$

Using the variables defined above for Qdot-labeled XPA, D_{lim} = 1.54 × 10^{-2} μm^2/s.

The energy barriers to free diffusion (E_A) can be calculated using the Arrhenius relationship:

$$k = e^{-E_A/k_BT} \quad \textbf{Equation 10}$$

where *k* is the rate constant (in this case, the stepping rate $2D/BP^2$) and E_A is the activation energy of the reaction. The energy barrier to free diffusion is the difference between the theoretical ("barrier-less") E_A and the experimentally determined E_A. By rearranging Equation 10 and substituting for *k*, this difference can be calculated:

$$\Delta E_A = ln\left(\frac{D_{lim}}{D_{expt}}\right) \times k_BT \quad \textbf{Equation 11}$$

where D_{expt} is the experimentally determined *D*. Because the ΔE_A can also be used to describe the roughness of the energy landscape, this value may also be referred to as σ (ref. [5]). XPA undergoing short-range linear diffusion had a mean *D* of 2.49 × 10^{-3} μm^2/s (Figure 13b), thus permitting the calculation of the energy barrier to diffusion via Equation 11, ΔE_A = 1.57 × k_BT. The diffusion coefficient for XPA the long-range mode (3.67 × 10^{-2} μm^2/s) exceeds the theoretical limit of a protein diffusing along the contour of DNA (Figure 13b), so the relevant energy barrier could not be calculated.

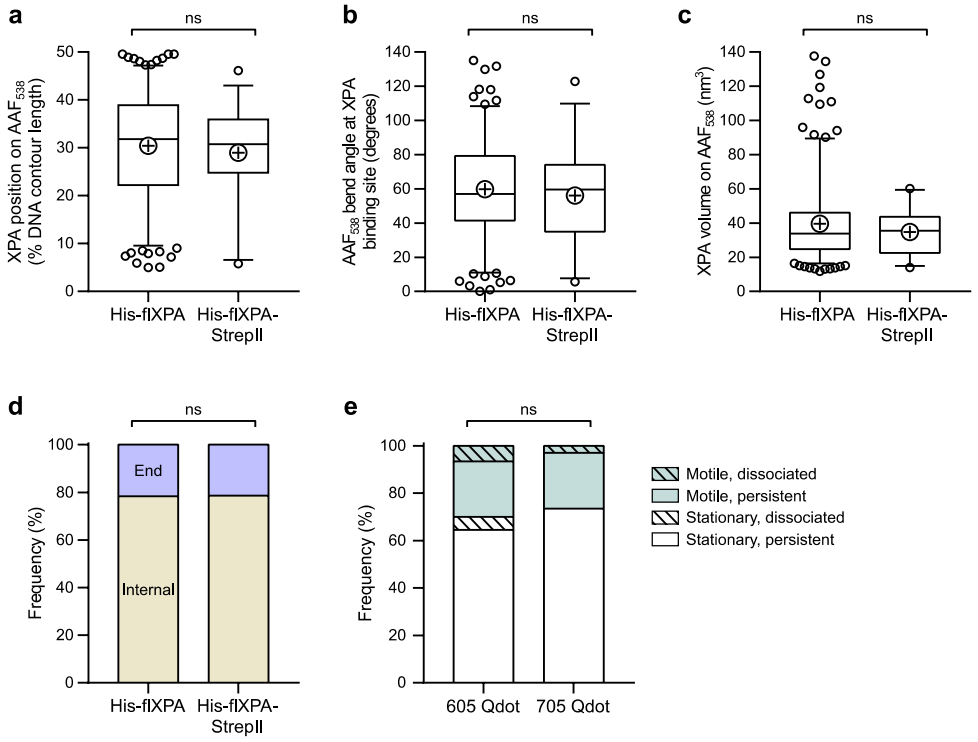

Figure 3. Comparison of XPA preparations and Qdot labeling strategies.

[a] Box and whisker plot (5-95 percentile) showing binding position on AAF_{538} of internally-bound His-XPA (n = 217, data reproduced from Figure 5e for comparison) and His-XPA-StrepII (n = 33). ns, p = 0.4810 by two-tailed Student's *t* test (p = 0.4628 by F test to compare variances). [b] Box and whisker plot (5-95 percentile) showing DNA bend angles at all sites of internally bound XPA on AAF_{538}. Results obtained with His-flXPA (n = 181, data reproduced from Figure 11h for comparison) and His-flXPA-StrepII (n = 32) are shown. ns, p = 0.4996 by two-tailed Student's *t* test (p = 0.8013 by F test to compare variances). [c] Box and whisker plot (5-95 percentile) showing AFM volumes of His-flXPA (n = 235, data reproduced from Figure 10d) and His-flXPA-StrepII (n = 35) on AAF_{538}. ns, p = 0.2289 by two-tailed Student's *t* test (p = 0.0001 by F test to compare variances). [d] Percentage of His-flXPA (n = 277, data reproduced from Figure 5c for comparison) and His-flXPA-StrepII (n = 42) bound to DNA at ends (lavender) or internally (tan) on AAF_{538}. ns, p = 0.9728 by χ^2 test. [e] Stacked bar graph showing the fraction of motile (teal) vs. stationary (white) and persistent (solid) vs. dissociating (diagonal lines) His-flXPA particles on $UV\lambda_{20J}$ tightropes. Results obtained with 605 Qdot (n = 107) and 705 Qdot (n = 34) labeling strategies are shown (see Section 2.6.2 Protein Labeling). Data reproduced as a sub-set of Figure 12b. ns, p = 0.4214 by χ^2 test.

3.0 Results

3.1 XPA Binds Specifically to a dG-C8-AAF Lesion

We first confirmed that XPA recognizes AAF by electrophoretic mobility shift assay (EMSA) as reported by others[85,90,98,99]. We generated binding isotherms of XPA by incubating a 37 bp DNA duplex (8 nM) with or without a single dG-C8-AAF adduct (AAF_{37} and NDF_{37}, respectively) and increasing amounts of purified full-length human XPA (His-flXPA-StrepII, Figure 4a). The apparent equilibrium dissociation constant (K_D) was 253.3 ± 14.2 nM for ND_{37} and 109.0 ± 4.6 nM for AAF_{37}, an approximately 2.3-fold difference (Figure 4d,f). Because AAF_{37} contains only one specific site among 36 undamaged bp, this difference can be multiplied by a factor of 37 to account for non-specific binding to the AAF_{37} substrate[3,46,94]. This results in approximately 85-fold specificity for dG-C8-AAF over non-damaged DNA, in good agreement with a previously reported specificity of XPA for 6-4PP[46]. In this way, we also found the specificity for a CPD lesion to be ~44-fold (Figure 4h). Furthermore, at higher XPA concentrations, a second band of higher molecular weight appeared, presumably indicating binding of a second XPA protein. It is important to note that any affinity of XPA for DNA ends could obscure EMSA results in terms of (a) specificity, as end-binding would increase overall binding on both substrates, thereby lowering the apparent specificity for the lesion, and (b) stoichiometry, as separate XPA proteins bound to the lesion and the end of the DNA would migrate the same as a true dimer in the gel.

We therefore turned to AFM to study XPA binding to a 538 bp DNA substrate with or without a single dG-C8-AAF lesion. The small size of XPA presents challenges in terms of

resolution and ability to visualize XPA bound to DNA using AFM in tapping mode; we thus adopted the use of PeakForce Tapping mode (Bruker) with a 2 nm tip to achieve improved resolution and reduced sample deformation[155-157]. Using this method, we were able to clearly recognize XPA bound to DNA by the increased AFM height and width of the complex.

Full-length human XPA (His-flXPA, Figure 4a, 1-4 μM) was incubated with a non-damaged 538 bp DNA substrate (ND_{538}, 100 nM in dsDNA fragments) and three-dimensional images were obtained using PeakForce Tapping AFM (Figure 5a). Of all the complexes observed, 33% were bound to the ends of the DNA substrate (Figure 5c). For the remaining internally-bound proteins, position along the DNA molecule was measured as a percentage of the total contour length of the DNA. XPA position revealed no preference for a specific internal site (Figure 5d).

XPA (His-flXPA) was then incubated with a 538 bp DNA substrate of the same sequence as ND_{538} but with a single dG-C8-AAF lesion at 30% from the 3' end (AAF_{538}, Figure 5b). On this substrate, only 22% of bound XPA proteins were found at the DNA ends. We also observed an increased frequency of complexes found near the lesion, at the expense of end-binders and other non-specific complexes (Figure 5e). A Gaussian was fit to the distribution of binding positions with mean 32.8 ± 12.3%. This spread of values is similar to others we have reported for lesion-binding proteins[22]. As published by Erie and colleagues, a Gaussian model with an additional term accounting for non-specific binding can be fit to position data obtained from AFM experiments and used to assess specificity without confounding end-binders[149]. Following their calculations, we find that XPA has a specificity for dG-C8-AAF of about 660 (see Section 2.5.3.3, Figure 6). Based on these data, it is clear that XPA is able to bind non-specifically to DNA (at ends and non-damaged sequences). However, the protein does exhibit specificity for the AAF adduct and binds preferentially at such a site.

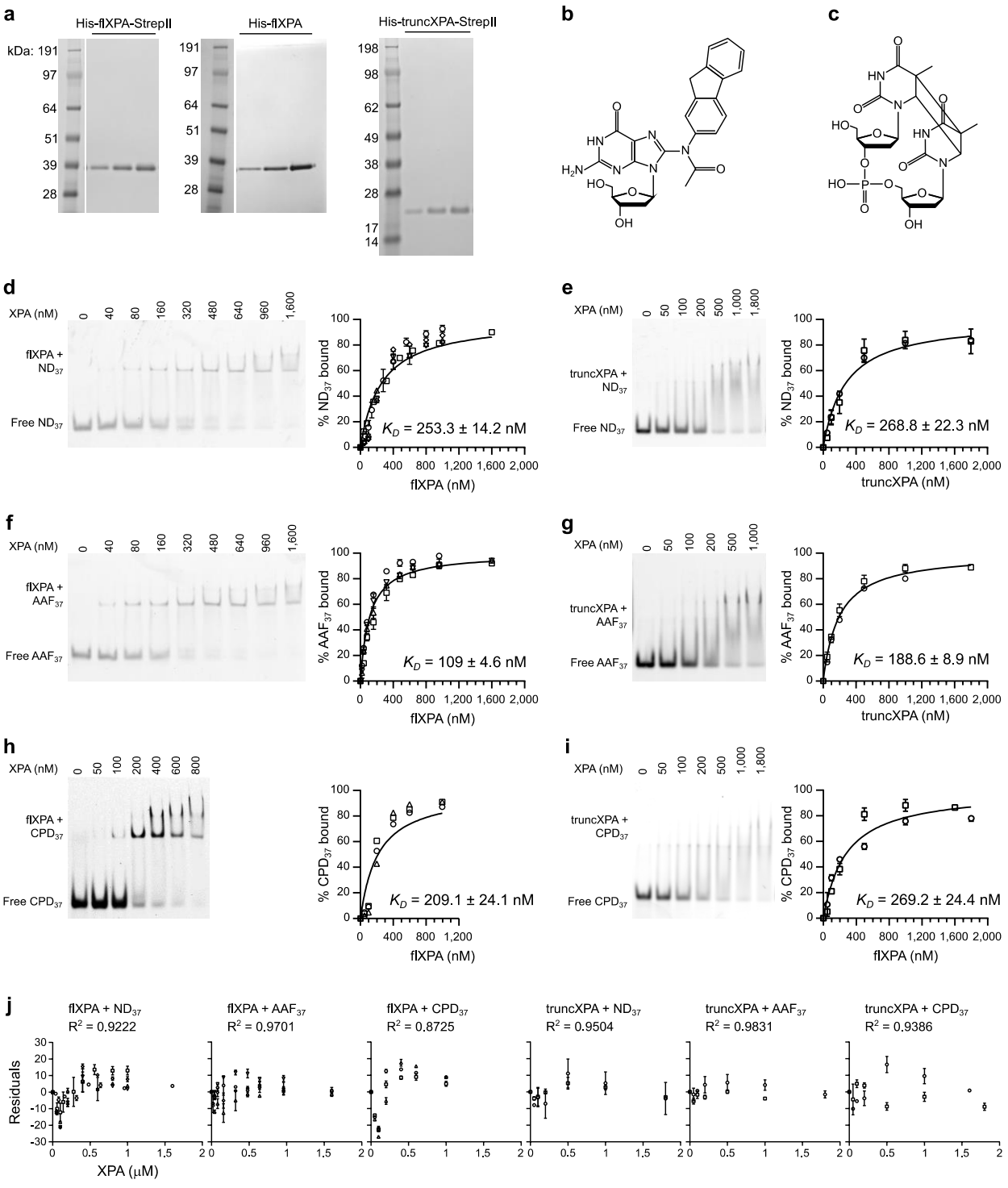

Figure 4. Purification and DNA binding activity of human XPA.

[a] SDS-PAGE and Coomassie stain of purified full-length XPA (flXPA) and a truncated variant containing residues M98 through T239 (truncXPA). Left, His-flXPA-StrepII (loading amounts: 270 ng, 540 ng, 1.08 μg).

Center, His-flXPA (160 ng, 490 ng, 1.14 μg). Right, His-truncXPA-StrepII (110 ng, 230 ng, 340 ng). Ladders shown from same gel with irrelevant lanes cut out. [b] Chemical structure of N-(2'-deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-C8-AAF). [c] Chemical structure of a cyclobutane pyrimidine dimer (CPD). [d] Left, representative EMSA gel showing flXPA binding to a 37 bp DNA substrate with a 5' fluorescein label (ND_{37}). Right, quantification of five experimental repeats (each run on duplicate gels) plotted as mean ± range. The equilibrium dissociation constant (K_D) was determined by a global fit to the data (see Section 2.4 for model) and is reported as best fit value ± s.e. of the fit. [e] Left, representative EMSA gel showing truncXPA binding to ND_{37}. Right, quantification of two experiments (run on duplicate gels) plotted/fit as in d. [f] Left, representative EMSA gel showing flXPA binding to a 37 bp DNA substrate with a central dG-C8-AAF adduct and a 5' fluorescein label (AAF_{37}). Right, quantification of four experimental repeats (each run on duplicate gels) plotted/fit as in d. [g] Left, representative EMSA gel showing truncXPA binding to AAF_{37}. Right, quantification of two experiments (run on duplicate gels) plotted/fit as in d. [h] Left, representative EMSA gel showing flXPA binding to a 37 bp DNA substrate with a central CPD lesion and a 5' fluorescein label (CPD_{37}). Right, quantification of three experimental repeats (each run on duplicate gels) plotted/fit as in d. [i] Left, representative EMSA gel showing truncXPA binding to CPD_{37}. Right, quantification of two experiments (run on duplicate gels) plotted/fit as in d. [j] Residuals and R^2 values for binding isotherm fittings of EMSA experiments, panels d–i. Residuals are plotted as mean ± range for each experiment (run on duplicate gels), against XPA concentration on the x-axis.

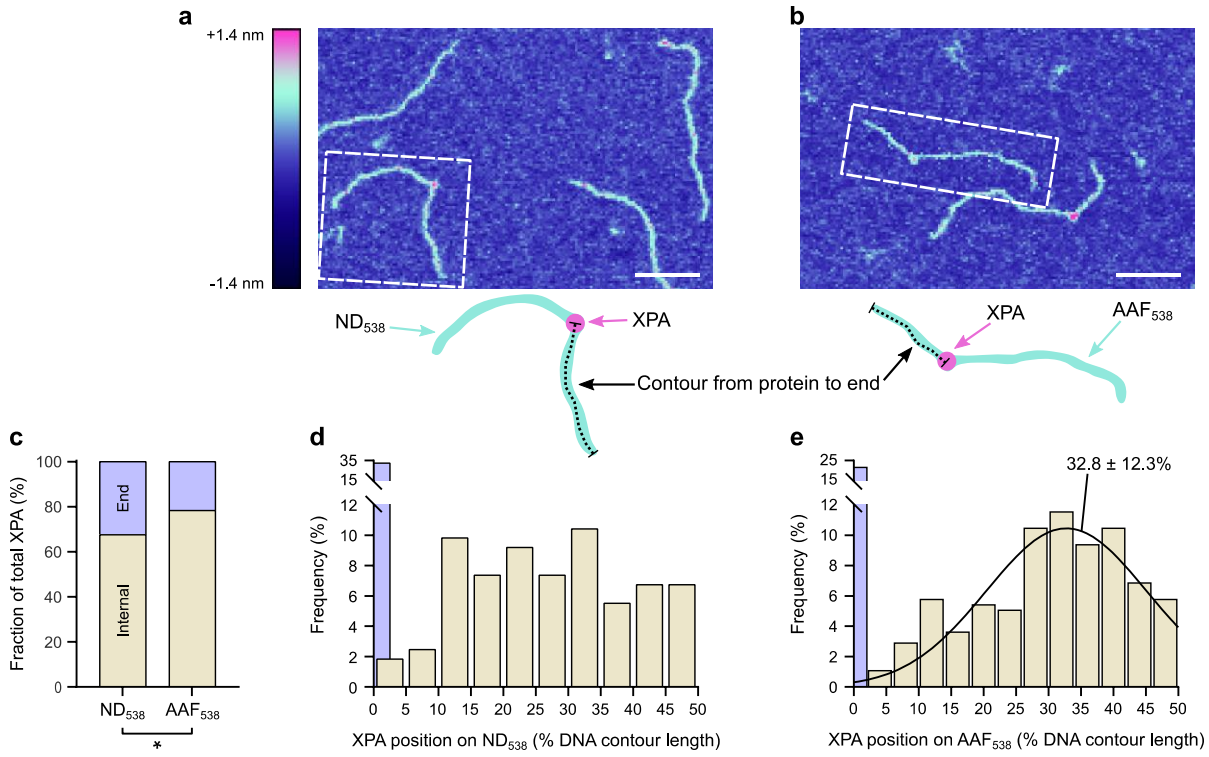

Figure 5. XPA binds specifically to a dG-C8-AAF lesion.

[a-b] Representative AFM image of XPA bound to a non-damaged 538 bp DNA substrate (a, ND_{538}) or a 538 bp DNA with a single dG-C8-AAF lesion at 30% from one end (b, AAF_{538}). Color scale represents AFM height and applies to panels a-b. White scale bar, 50 nm. The dashed white line indicates the example in the cartoon below. Binding position was measured between the center of the protein to the closest DNA end as a percentage of total DNA contour length. [c] Percentage of XPA bound to DNA at ends (lavender) or internally (tan) on ND_{538} (n = 163 particles) and AAF_{538} (n = 277 particles). * p = 0.0118 by χ^2 test. [d] Histogram showing the distribution of internally bound XPA (n = 110 particles) position on ND_{538}. End-binders are shown in lavender. [e] Histogram and Gaussian fitting of internally bound XPA (n = 217 particles) position on AAF_{538}. End binders are shown in lavender (not included in Gaussian fit). Gaussian is labeled with mean and s.d.

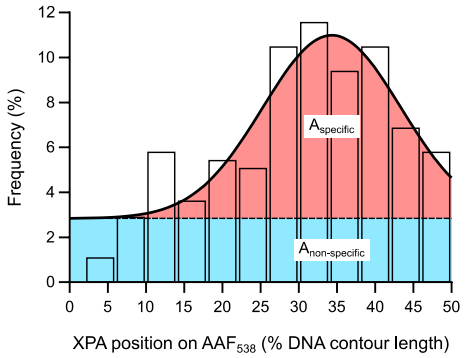

Figure 6. Specificity analysis of XPA binding position by AFM.

Histogram showing distribution of internally bound XPA on AAF_{538} (n = 217 particles, data reproduced from Figure 5e). To calculate specificity, a Gaussian model with an additional term for non-specific binding was fit to the data. Red ($A_{specific}$), area under curve representing specific binding. Blue ($A_{non\text{-}specific}$), area under curve representing non-specific binding. See Section 2.5.3.3 for calculation details.

3.2 XPA is a Monomer in Solution

To resolve the question of XPA stoichiometry, we first sought to clarify the oligomeric status of the free protein. While there is support that XPA is a monomer in solution[66,158,159], there have also been reports that it forms dimers and higher oligomers[89]. AFM has been successfully used to determine protein stoichiometry due to the linear relationship between AFM volumes of globular proteins and their molecular weight[160,161]. Full-length human XPA (His-flXPA, 32.6 kDa) was diluted to 40 nM and deposited on mica for imaging by PeakForce Tapping AFM in air. Volumes of the particles were measured and the data fit a Gaussian distribution centered at 30.3 ± 15.4 nm^3 (Figure 7a).

In order to translate this volume into molecular weight, and thus protein stoichiometry, we generated a standard curve using monomeric proteins of known sizes (Figure 7b). HMGB1 (25

kDa), APE1 (37 kDa with His tag), DNA polymerase β (Polβ, 42.8 kDa with His tag), and UvrD (85.6 kDa with His tag) were adsorbed at 40 nM on mica and imaged by PeakForce Tapping AFM (Figure 8). Measured AFM volumes of these proteins were plotted against their known molecular weight and fit using least-squares linear regression: $Volume = (MW \times 1.14) - 2.00$. Using this equation, the measured volumes for XPA correspond to 28.4 ± 15.3 kDa, close to the expected molecular weight of 32.6 kDa for the monomer. There was no significant population of XPA corresponding to the dimer size. These data suggest that, at 40 nM in solution, XPA exists as a monomer.

To confirm that XPA is a monomer at higher protein concentrations, we performed size exclusion chromatography coupled with multiangle light scattering (SEC-MALS). At 65 μM and 80 μM, purified XPA (His-flXPA-StrepII, 33.9 kDa) eluted in a single major peak, with a molecular weight corresponding principally to a monomer (Figure 7c). By both methods, the apparent molecular weight was a few kDa less than the theoretical value based on the sequence with tags, potentially due to protein conformation. Taken together, these data clearly indicate that XPA predominantly exists as a monomeric species in solution, in agreement with previous reports[159,162].

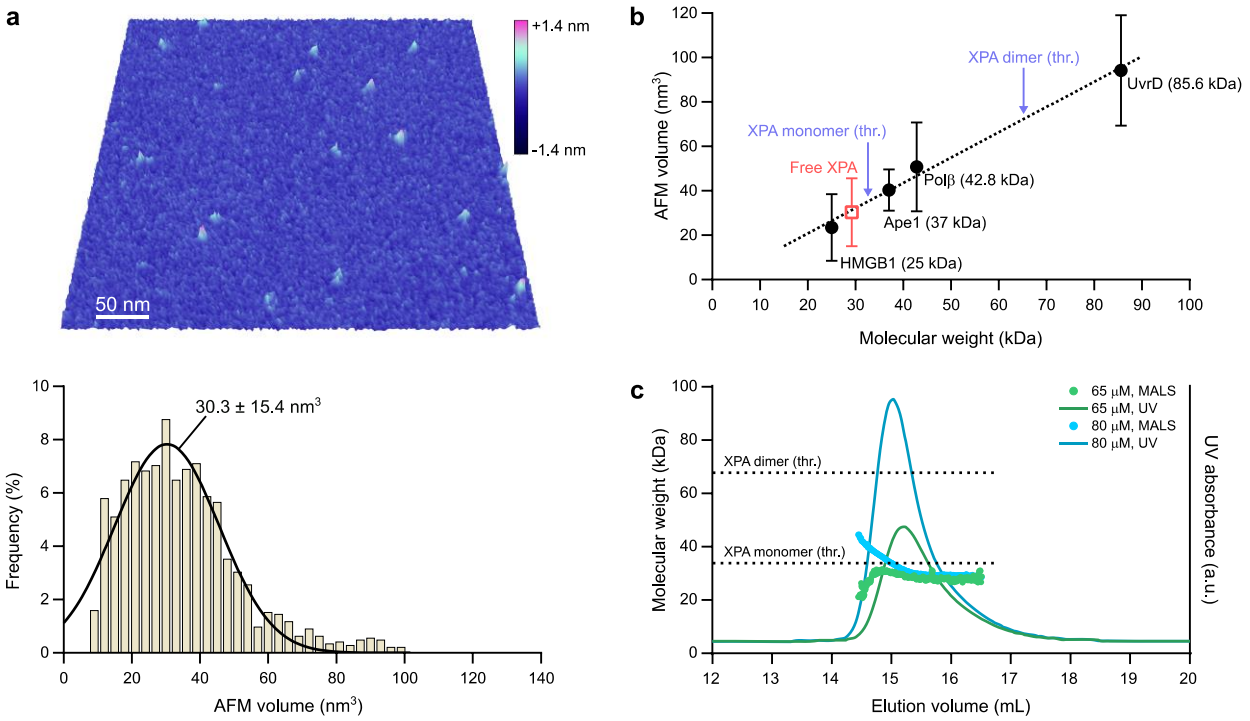

Figure 7. XPA is a monomer in solution.

[a] Top, representative 3D AFM image of free XPA. Color scale represents AFM height. Bottom, histogram and Gaussian fit of free XPA proteins imaged by AFM (n = 1,451 particles). Gaussian is labeled with mean and s.d. [b] Standard for calculating molecular weight from AFM volumes. Volumes of free proteins imaged by PeakForce Tapping AFM were calculated and respective histograms were fit by Gaussian distributions. Solid black circles, proteins used to generate the standard plotted against known MW (results and n values for each in Figure 8). Dashed line, linear regression, resulting in the calibration curve: Volume (nm^3) = 1.14 × MW (kDa) – 2.00. R^2 = 0.990. Purple arrows point to theoretical (thr.) volumes for the purified His-flXPA monomer (32.6 kDa) and dimer (65.2 kDa). Open red square, experimental His-flXPA AFM volume (see a), corresponding to a molecular weight of 28.4 ± 15.3 kDa. Errors bars indicate s.d. of the Gaussian distribution. [c] Molar mass determination by SEC-MALS of XPA at 65 μM (green) and 80 μM (blue). Theoretical molecular weights for the purified His-flXPA-StrepII monomer (33.9 kDa) and dimer (67.8 kDa) are indicated by dashed lines

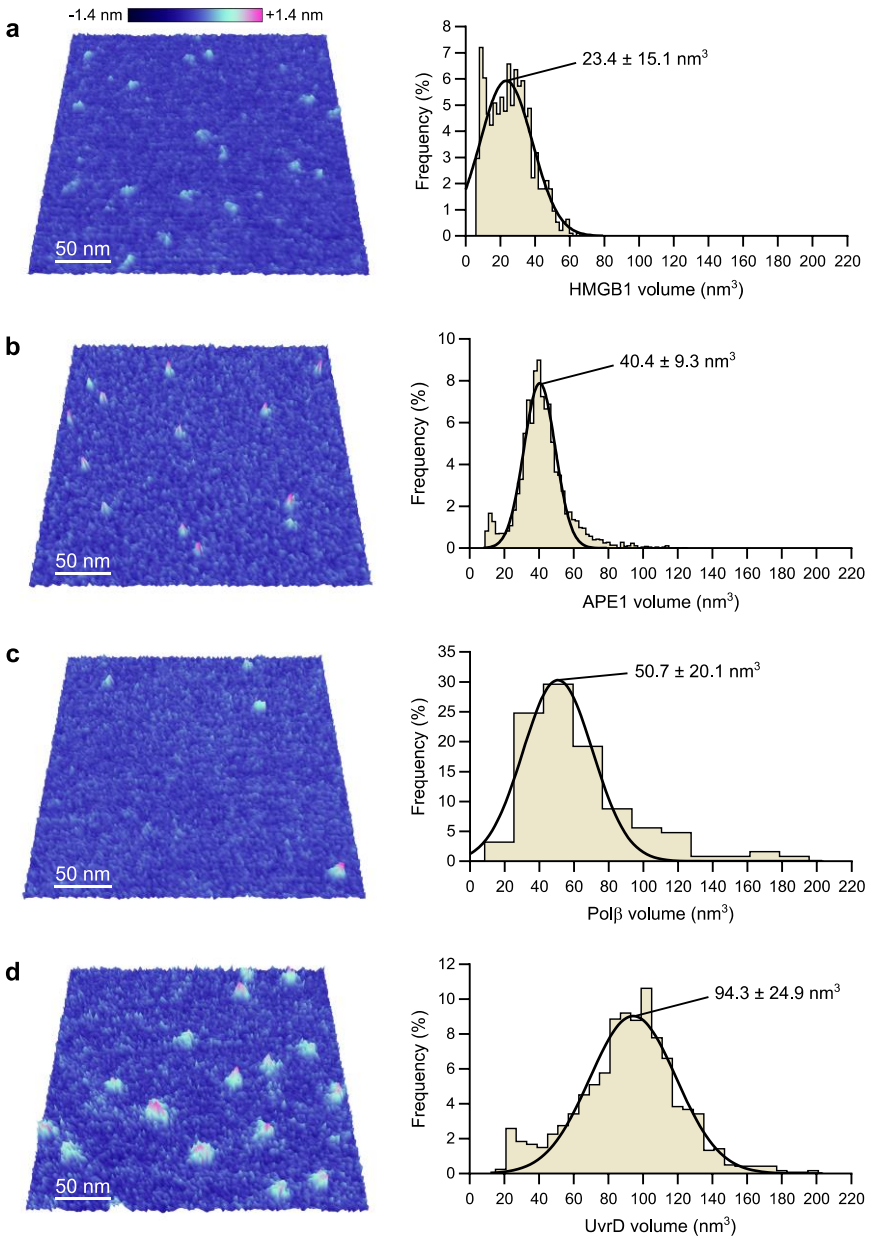

Figure 8. Generation of standard for AFM volumes.

[a-d] Left, representative 3D AFM image and right, histogram and Gaussian fitting of AFM volumes used to generate the standard curve shown in Figure 7b. Gaussians are labeled with mean ± s.d. [a] HMGB1, 25 kDa. n = 943. Color scale represents AFM height and applies to all panels. White scale bar, 50 nm. [b] APE1, 37 kDa. n = 3,529. [c] Polβ, 42.8 kDa. n = 125. [d] UvrD, 85.6 kDa. n = 1,195.

3.3 XPA Binds Non-Damaged DNA and dG-C8-AAF as a Monomer

Previous reports suggest that XPA may bind DNA as a homodimer[89,90] while others have indicated monomeric binding[68]. AFM offers the unique ability to clearly distinguish between distinct complexes (e.g. one protein at a DNA end and one protein at a lesion) and true dimerization. Therefore, we measured the AFM volumes of XPA bound to DNA in order to determine stoichiometry. Because XPA is a small protein, we expected the DNA to contribute a significant amount of volume to the total complex. Some groups have considered this issue and have reported AFM volumes with the DNA portion subtracted[163]. Based on this idea, we developed a method to determine the size of the DNA within the complex and subtract its volume to obtain the volume of XPA alone (Figure 9a).

It was important to validate that we would be able to use the standard curve based on free proteins (Figure 7b) to analyze the volumes of bound proteins. Using two proteins of similar size and known stoichiometry and the methods described, we were able to successfully determine molecular weights based on the standard. APE1 (37 kDa) was incubated with a 514 bp DNA substrate containing a nick at 36% from one end ($Nick_{514}$) and imaged by AFM. The distribution of AFM volumes for APE1 on the DNA was centered at 40.6 ± 10.7 nm^3 (Figure 9c). This is very close to the AFM volume obtained for the free protein (40.4 nm^3, Figure 8b). Furthermore, using the standard curve, this corresponds to a molecular weight of 37.4 ± 11.2 kDa. Using our methods for DNA volume subtraction combined with the standard of free proteins, we were able to accurately estimate the molecular weight of APE1. We repeated this test using Polβ (42.8 kDa). Again, the protein was incubated with $Nick_{514}$ and imaged by AFM. The distribution of AFM volumes was centered at 31.9 ± 16.1 nm^3, which corresponds to 29.8 ± 15.9 kDa (Figure 9e) and was smaller than expected. The AFM volume of the free protein was 50.7 nm^3 (Figure 8c),

suggesting that, in this case, we have over-estimated the contribution of the DNA to the total complex volume. This is a very likely explanation if the crystal structures of APE1 and Polβ are taken into consideration (insets, Figure 9c,e). APE1 is positioned on top of the DNA such that the volume of the complex is essentially the sum of the protein and the DNA alone, as our model assumes. However, Polβ pulls apart the nicked DNA backbone and inserts itself much further into the helix[164]; in this case, the volume of the complex appears to be less than the sum of the two parts. Overall, we can conclude that both APE1 and Polβ bind DNA as a monomer using our method, but it is important to note that the accuracy of molecular weight estimates is dependent on the precise conformation of the protein-DNA complex.

Having confirmed that we would be able to distinguish XPA monomers and dimers on DNA, we measured AFM volumes for all XPA proteins (His-flXPA, 32.6 kDa) bound to the ND_{538} and AAF_{538} substrates at multiple concentrations as high as 16-fold above the K_D of XPA for damaged and non-damaged DNA (see Section 2.5.3.3, Figure 10a-b). In some cases, we observed multiple binding events on the same DNA molecule; these were measured individually if there was a clear stretch of unbound DNA between them. XPA bound to ND_{538} had a distribution of AFM volumes centered at 28.4 ± 12.7 nm^3, corresponding to 26.7 ± 12.9 kDa (Figure 10c). The volumes of internally and end-binding proteins had similar distributions. XPA bound to AAF_{538} had a distribution of volumes centered at 31.0 ± 12.9 nm^3, corresponding to 29.0 ± 13.1 kDa (Figure 10d). Again, the volumes of internally- (either near the lesion or not) and end-binding proteins had similar distributions. These data suggest that, regardless of where XPA binds DNA, it does so as a monomer, which is consistent with an earlier study[68]. We did not observe a significant secondary population corresponding to the dimer.

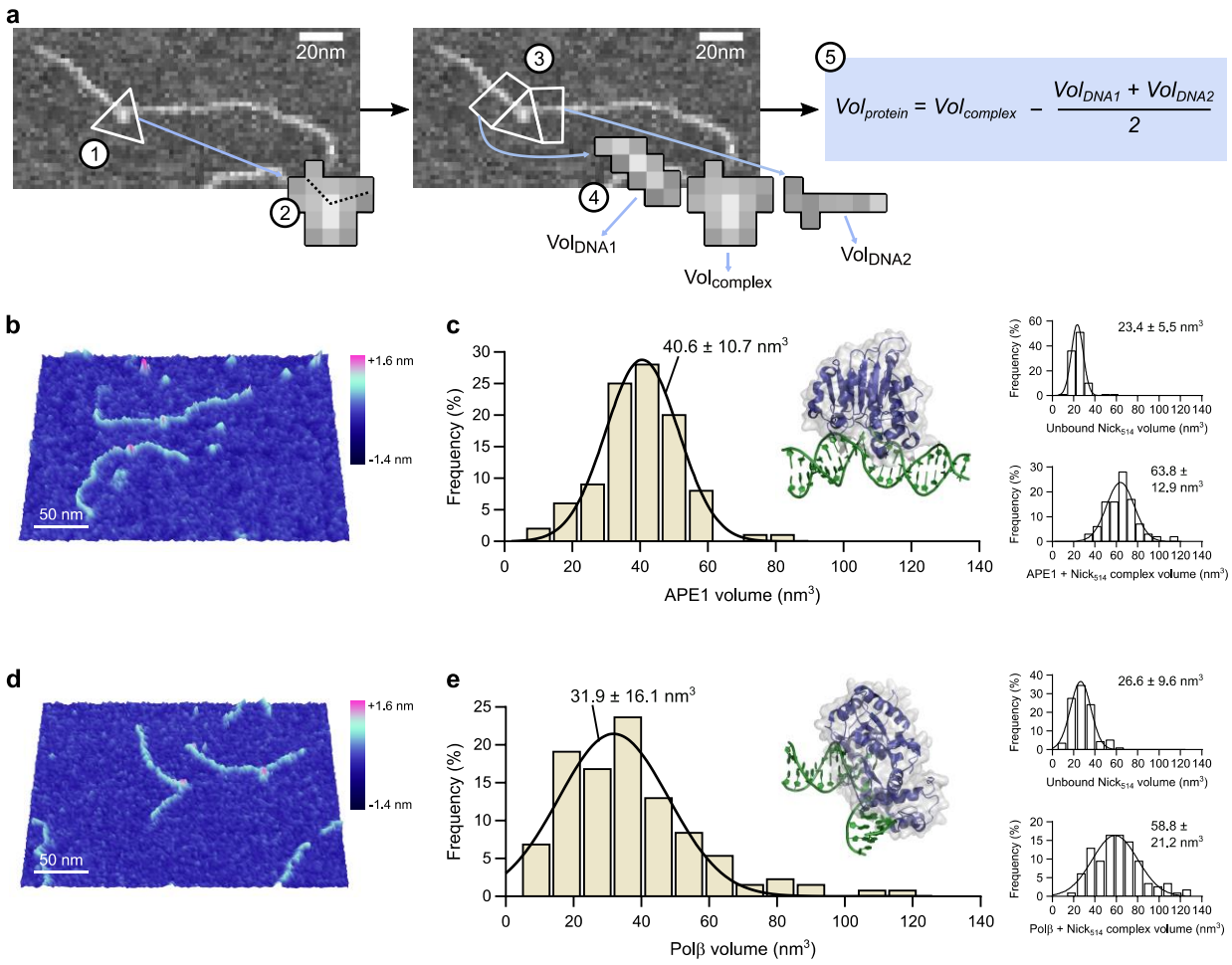

Figure 9. Determination of DNA-bound protein AFM volume.

[a] Schematic showing the steps used to subtract DNA volume from a protein-DNA complex. (1) Outline perimeter of the complex to separate it from unbound DNA. (2) Measure contour length of the DNA path through the complex. (3) Delineate regions of unbound DNA on either side of the complex with the same length measured in step 2. (4) Obtain AFM volumes for all three regions. (5) Protein volume is calculated as the volume of the complex minus the average DNA volume. [b] Representative 3D AFM image of APE1 bound to a 514 bp DNA substrate with a nick at 36% from one end ($Nick_{514}$). [c] Left, histogram and Gaussian fitting of the distribution of calculated AFM volumes of APE1 on $Nick_{514}$ (n = 100). The Gaussian is labeled as mean ± s.d. The AFM volume corresponds to 37.4 ± 11.2 kDa. Inset, crystal structure of APE1 bound to a nicked abasic DNA substrate (PDB 5DFF). Right, histograms and Gaussian fittings of measured volumes of unbound DNA and total complex. [d] Representative 3D AFM image of Polβ bound to $Nick_{514}$. [e] Left, histogram and Gaussian fitting of the distribution of calculated AFM volumes (n = 131). The Gaussian is labeled as mean ± s.d. AFM

volume corresponds to 29.8 ± 15.9 kDa. Inset, crystal structure of Polβ bound to nicked DNA (PDF 1BPZ). Please see text for discussion of the underestimation of the protein size. Right, histograms and Gaussian fittings of measured volumes of unbound DNA and total complex

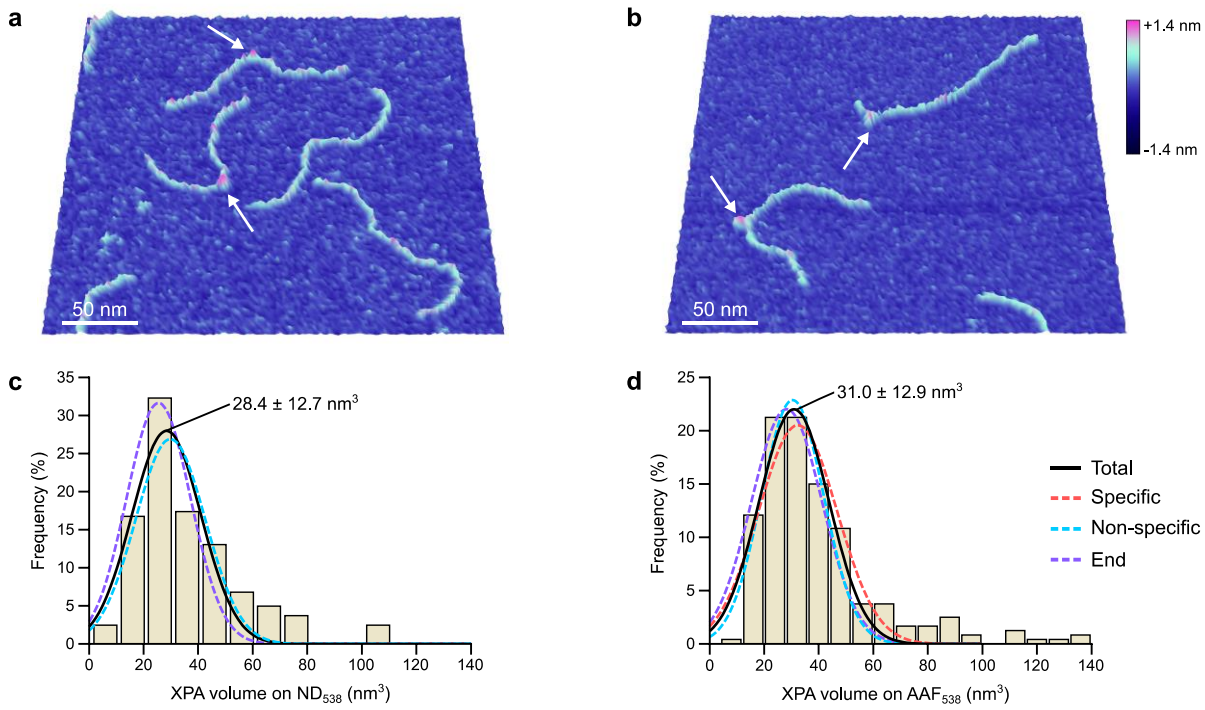

Figure 10. XPA binds non-damaged DNA and dG-C8-AAF modified DNA as a monomer.

[a-b] Representative 3D AFM image of XPA bound to 538 bp non-damaged DNA (a, ND_{538}) or 538 bp DNA with a single dG-C8-AAF adduct at 30% from one end (b, AAF_{538}). White arrows point to XPA bound to DNA. Color scale represents AFM height and applies to panels a-b. [c] Histogram showing distribution of AFM volumes of all XPA proteins bound to ND_{538} (n = 161 particles) with Gaussian fit (solid black line). AFM volume corresponds to 26.7 ± 12.9 kDa, using the standard shown in Figure 7b. Dashed cyan line, Gaussian fit to the sub-fraction of non-specifically bound XPA (i.e. all internal complexes, n = 108); centered at 29.9 nm^3 (s.d. 12.8 nm^3), corresponding to 28.0 ± 13.0 kDa. Dashed purple line, Gaussian fit to the sub-fraction of XPA bound at DNA ends (n = 53); centered at 25.5 nm^3 (s.d. 11.7 nm^3), corresponding to 24.1 ± 12.0 kDa. [d] Histogram showing the distribution of AFM volumes of all XPA proteins (n = 235 particles) bound to AAF_{538} with a Gaussian fit (solid black line). The AFM volume corresponds to 29.0 ± 13.1 kDa. Dashed red line, Gaussian fit to the sub-fraction of XPA bound between 20 and 40% of the DNA contour length ("specific," n = 58); centered at 32.4 nm^3 (s.d. 14.4 nm^3), corresponding to 30.2 ± 14.4 kDa. Dashed cyan line, Gaussian fit to the sub-fraction of XPA bound internally but at positions away from the lesion ("non-specific," n = 84); centered at 30.5 nm^3 (s.d. 11.4 nm^3), corresponding to 28.5 ± 11.8 kDa. Dashed purple line, Gaussian fit to the sub-fraction of XPA bound at DNA ends (n = 51); centered at 28.3 nm^3 (s.d. 13.0 nm^3), corresponding to 26.6 ± 13.2 kDa.

3.4 XPA Bends DNA ~60°

Having shown that XPA binds to damaged and non-damaged DNA as a monomer, we next asked whether XPA induces a bend at the dG-C8-AAF site. We first sought to determine if the dG-C8-AAF lesion itself introduces any flexibility into the DNA helix. Because this lesion was placed at 30% from the 3' end of the AAF_{538} substrate, we measured the DNA bend angle at this site for unbound DNA. Note that while there is only a single lesion in our damaged substrate, we measured the angle at 30% from both ends (resulting in two angles per DNA molecule) because, under the current conditions, we are unable to differentiate between the 5' and 3' end of the molecule. As a negative control we measured the DNA angle at 30% from each end on the 538 bp non-damaged substrate, ND_{538} (Figure 11a), and found the majority of angles were around 0° (Figure 11c). The distribution for the dG-C8-AAF-modified DNA, AAF_{538}, showed the emergence of a second population of bend angles (Figure 11b,d). A double Gaussian fit to the data describes two populations at 10.5 ± 7.0° and 34.8 ± 10.6°. While the smaller angle likely represents the unmodified site, the larger angle is likely the result of flexibility introduced by dG-C8-AAF[165,166].

We next sought to determine if XPA bends the DNA. By AFM, we see that XPA bends both non-damaged (Figure 11e) and AAF-adducted (Figure 11f) DNA. DNA angles were measured at all sites of internally-bound XPA. On ND_{538}, XPA induced a bend angle of 54.0 ± 30.1° (Figure 11g). On AAF_{538}, XPA induced a bend angle of 58.6 ± 26.8°; the distribution was essentially the same for XPA bound near the lesion or non-specifically (Figure 11h). This angle is greater than that introduced by the lesion itself (~30°). Together, these data indicate that XPA binds preferentially to the helix-bending dG-C8-AAF and that, regardless of binding site, the complex contains a single XPA protein and the DNA is bent ~60°.

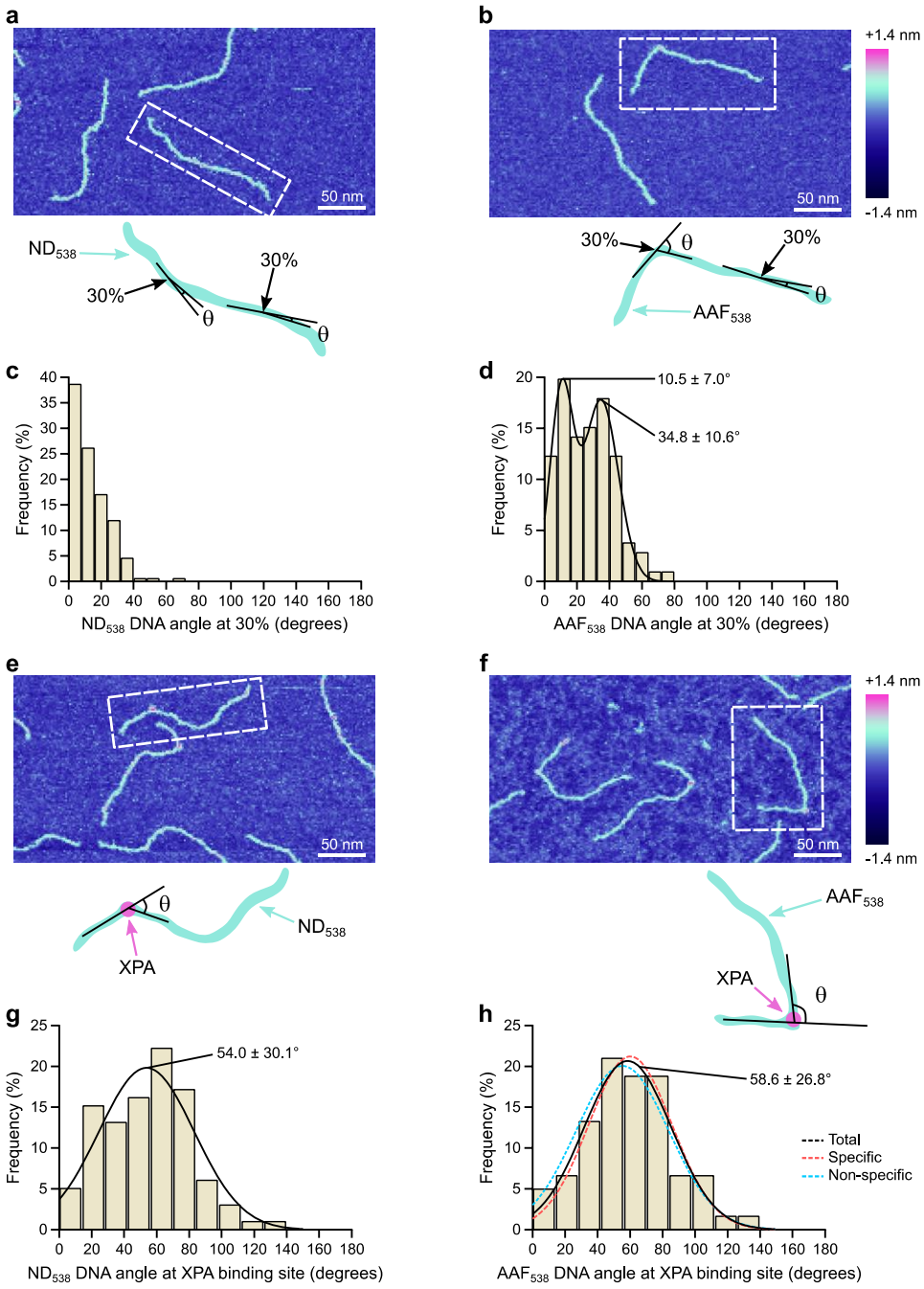

Figure 11. dG-C8-AAF introduces DNA bending and XPA bends DNA ~60° when it binds.

[a-b] Representative AFM image of free 538 bp non-damaged DNA (a, ND_{538}) or free 538 bp DNA with a single dG-C8-AAF adduct at 30% from one end (b, AAF_{538}). Color scale represents AFM height and applies to all panels. Dashed white line indicates the example in the cartoon below. Angles were measured at sites 30% of the total DNA contour length from each end and are reported as the θ angle (supplement to internal DNA angle).

[c] Histogram and Gaussian fitting of the inherent DNA bend angle of non-damaged DNA (n = 176 angles). [d] Histogram and double Gaussian fitting of the inherent DNA bend angle of AAF-adducted DNA (n = 106 angles). Each peak of the Gaussian is labeled as mean ± s.d. [e-f] Representative AFM image of XPA bound to ND_{538} (e) or AAF_{538} (f). The dashed white line indicates the example in the cartoon below. Angles were measured at sites of bound XPA and are reported as the θ angle. [g] Histogram showing the distribution of DNA bend angles at all sites of internally bound XPA on ND_{538} (n = 99 angles). [h] Histogram showing distribution of DNA bend angles at all sites of internally bound XPA on AAF_{538} (n = 181 angles) with Gaussian fit (solid black line). Gaussian is labeled as mean ± s.d. Dashed red line, Gaussian fit to the sub-fraction of angles at XPA sites between 20 and 40% of the DNA contour length ("specific," n = 116); centered at 60.2° (s.d. 25.7°). Dashed cyan line, Gaussian fit to the sub-fraction of angles at sites of XPA bound internally but at positions away from the lesion ("non-specific," n = 65); centered at 55.3° (s.d. 28.5°).

3.5 XPA Performs Episodic One-Dimensional Diffusion to Search DNA for Damage

The width of the Gaussian distribution of XPA binding position on AAF_{538} (i.e. observed non-specific binding around the dG-C8-AAF site) suggests that the protein might be dynamic on DNA. We have previously shown that Rad4-Rad23, which adopts a similar distribution of binding positions by AFM, performs constrained linear diffusion around a lesion[22]. Therefore, we next sought to investigate whether XPA displays one-dimensional diffusion on non-damaged and damaged DNA using a single molecule DNA tightrope assay[138,140,142].

DNA tightropes consisted of long (> 40 kbp) DNA molecules suspended between poly-L-lysine-coated silica microspheres in a flow cell. To visualize XPA on these tightropes, the purified His-tagged protein was labeled with either a streptavidin-conjugated 705 nm quantum dot (Qdot) and biotinylated anti-His antibody (Figure 12a) or an anti-mouse IgG antibody-conjugated 605 nm Qdot and mouse anti-His antibody. Both labeling strategies resulted in similar observed behavior

(Figure 3e). For observation, flow was turned off and Qdots were excited with a 488 nm laser at an oblique angle (optimized to illuminate particles above the surface of the flow cell, in the plane of suspended DNA) and 300 second movies were recorded at 10-12 frames per second (fps) of XPA on one of three different tightrope substrates: non-damaged genomic λ DNA (NDλ), UV-treated λ DNA ($UV\lambda_{20J}$ and $UV\lambda_{80J}$), and defined arrays of dG-C8-AAF (AAF_{array}). We have previously determined that exposure of λ DNA to 20 J/m^2 254 nm UV radiation results in a lesion density of approximately one UV photoproduct every 2.2 kbp[21]. Therefore, λ DNA was treated with 20 J/m^2 or 80 J/m^2 (producing 4X lesion density, or approximately one lesion every 550 bp) 254 nm UV-C to generate $UV\lambda_{20J}$ and $UV\lambda_{80J}$, respectively. Because UV-C radiation leads to formation of lesions comprising ~75% CPDs and ~25% 6-4PPs[25], we expect $UV\lambda_{80J}$ to contain one 6-4PP every ~2.2 kbp. AAF arrays were prepared via end-to-end ligation of a 2030 bp fragment of linear DNA with a single site-specific dG-C8-AAF modification[21,142]. Kymographs were first categorized into four groups: stationary/persistent, stationary/dissociated, motile/persistent, and motile/dissociated. Motility was defined as linear displacement greater than 130 nm (three pixels, see Section 2.6.4 Data Analysis) over the course of observation. XPA was primarily stationary (60-70% of the molecules) on all tightrope substrates (Figure 12b, flXPA). Although the proportions of these broad categories do not appear to be affected by the DNA substrate, we expected the nature of XPA's motion to differ. Further detailed analyses of the motile fraction (i.e. those that moved at least once) are presented here and below.

Unlike other repair proteins we have observed at the single molecule level[21-23], it is interesting to note that the motile XPA particles often switched their behavior multiple times during observation (Figure 12c). Most exhibited some periods of pausing between episodic phases of linear diffusion. To analyze these differences, we categorized the behavior of motile XPA into

three distinct modes: paused (particle not moving), short-range diffusion (displacement less than 690 nm), and long-range diffusion (displacement greater than 690 nm). Thresholds for motile modes were chosen based on approximate distances between lesions; assuming true B-form DNA, 2,030 bp corresponds to 690 nm. Figure 12c shows an example of XPA on NDλ, exhibiting all three modes. As shown, the behavior of each particle was divided into phases based on these modes.

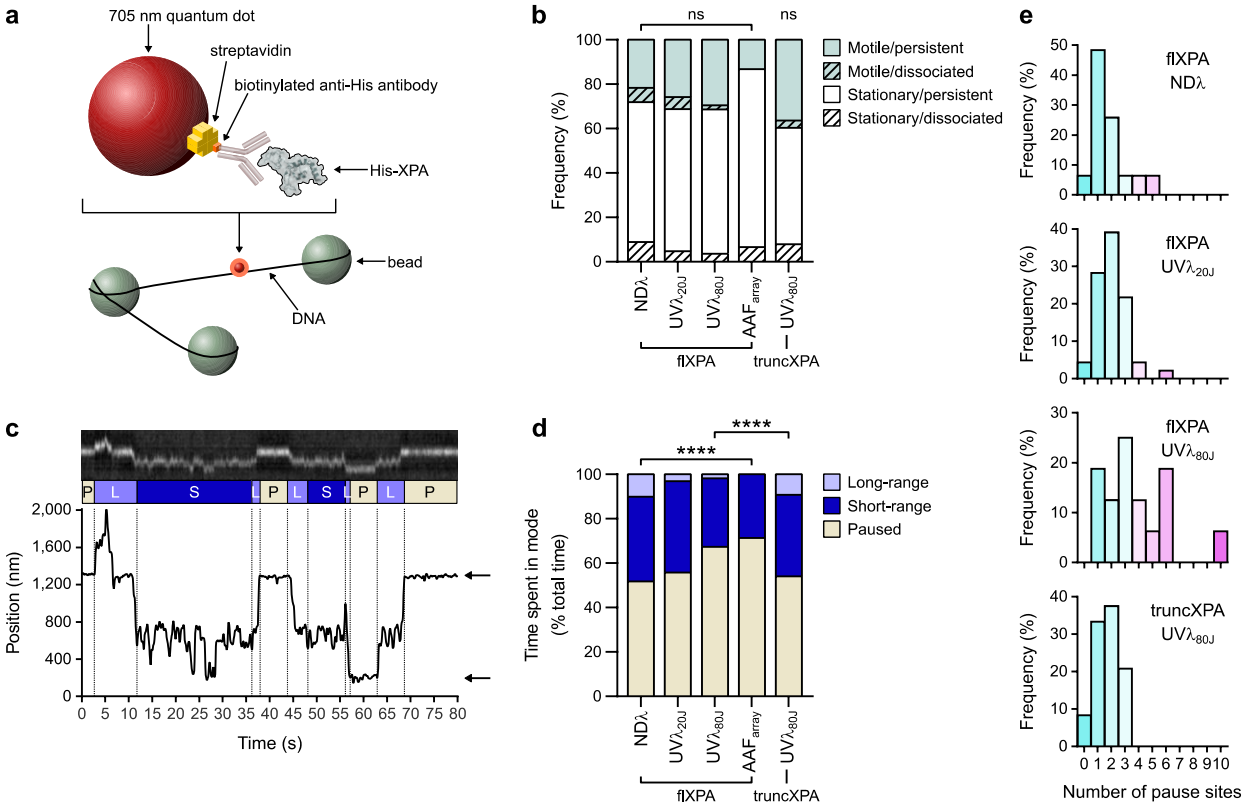

Figure 12. XPA exhibits episodic linear diffusion on DNA tightropes. DNA damage leads to increased pausing, dependent on N- and C-termini.

[a] Cartoon showing one strategy used for XPA labeling on DNA tightropes. His-tagged XPA is labeled with a biotinylated anti-His antibody bound to a streptavidin-conjugated 705 nm quantum dot. See Section 2.6.2 for alternative labeling strategy. [b] Stacked bar graph showing the fraction of motile (teal) vs. stationary (white) and persistent (solid) vs. dissociating (diagonal lines) particles of full-length XPA (flXPA) on non-damaged λ (NDλ, n = 124 particles), 20 J/m^2 UV-irradiated λ (UVλ$_{20J}$, n = 147), 80 J/m^2 UV-irradiated λ (UVλ$_{80J}$, n = 54), and AAF arrays (AAF$_{array}$, n = 45), and of truncated XPA (truncXPA) on UVλ$_{80J}$ (n = 63). Motile particles are defined as those which moved more than 130 nm on DNA during 300 s observation. ns, no significant difference between groups by χ^2 for all flXPA experiments, for all flXPA and truncXPA categories, or for truncXPA on UVλ$_{80J}$ vs. flXPA on UVλ$_{80J}$. [c] Example of a motile kymograph (cut to show only 80 s of the recorded movie) of 705 nm quantum dot-labeled XPA on NDλ. Particle position (bottom) was localized using Gaussian fittings to the intensity profile on the fluorescence image (top). Dashed lines separate phases and diffusive modes are labeled according to particle displacement: paused (P, displacement < 130 nm, tan), short-range motion (S,

displacement 130 – 690 nm, navy), and long-range motion (L, displacement > 690 nm, lavender). Arrows point to pause sites occupied by particle. [d] Stacked bar graph showing the fraction of time spent in each mode as percentage of total time recorded for all motile particles. ****, $p < 0.0001$ by χ^2 for all flXPA experiments and for truncXPA on $UV\lambda_{80J}$ vs. flXPA on $UV\lambda_{80J}$. [e] Histogram of number of pause sites for flXPA on NDλ (n = 31 particles), $UV\lambda_{20J}$ (n = 46), and $UV\lambda_{80J}$ (n = 16), and for truncXPA on $UV\lambda_{80J}$ (n = 24).

3.6 Presence of DNA Damage Increases Pausing in Motile XPA Particles

Notably, the presence of damage in the DNA tightrope had an impact on the occupancy of each mode. The proportion of seconds spent in each mode was calculated as a fraction of total recorded time for motile particles (Figure 12d). Full-length XPA (flXPA) demonstrated a clear dose-dependent increase in paused time with increasing damage. Motile particles of flXPA spent 52% of the time paused on NDλ tightropes; this increased to 56% on $UV\lambda_{20J}$, 67% on $UV\lambda_{80J}$, and 71% on AAF_{array}, and was accompanied by a decrease in time spent in the diffusive modes, especially long-range mode.

Another unique feature of XPA linear diffusion on DNA was that the protein appeared to prefer and return to certain positions on the tightropes. These positions, deemed "pause sites," were defined as sites at which XPA spent at least one paused phase (at least five seconds), and were examined as a second measure of particle pausing. On NDλ, the majority of XPA particles (48.4%) paused at just one position (Figure 12e, top row). On $UV\lambda_{20J}$, the majority of XPA particles (39.1%) paused at two distinct positions (Figure 12e, second row), and on $UV\lambda_{80J}$, the majority (25.0%) occupied three distinct positions, with some occupying as many as 10 positions (Figure 12e, third row). The increasing number of pause sites observed with increasing UV lesion density in the tightropes strongly suggest that these positions correspond to sites of damage. Pause

sites observed on non-damaged DNA might represent other distorted regions of the DNA (e.g. A-tract-GC junctions or spontaneous damage sites). Future work will be necessary to confirm the nature of these pause sites. Due to the small sample size of motile XPA particles on AAF_{array} (n=6), this dataset was not included in this and subsequent analyses.

3.7 Truncated XPA Exhibits Reduced Pausing on Damaged DNA

We next examined the behavior of a truncated XPA mutant (His-truncXPA-StrepII, or truncXPA) to investigate the conformation of XPA during linear diffusion and pausing. We hypothesized that the intrinsically disordered N- and C-terminal arms of XPA play a role during target search. Thus, the N- and C-termini were deleted, leaving only the currently accepted DNA-binding domain, M98 through T239 (ref. [77]). truncXPA maintains its ability to bind non-damaged 37 bp DNA (K_D = 268.8 ± 22.3 nM) with similar affinity as flXPA (Figure 4e). Specific binding of truncXPA to an AAF adduct (K_D = 188.6 ± 8.9 nM) is only ~53-fold (Figure 4g). Thus, we observed a reduction in specificity to dG-C8-AAF by a factor of 1.6 in the truncated XPA compared to full-length. All specificity was lost for CPD (K_D = 269.2 ± 24.4 nM, Figure 4i).

We turned to our DNA tightrope platform to test truncXPA target search on $UV\lambda_{80J}$ tightropes. Statistically, truncXPA exhibited the same proportion of stationary/motile and persistent/dissociated particles as flXPA (Figure 12b). Importantly, for motile particles, we observed a significant decrease (~25%) in time spent paused and a corresponding increase in time spent undergoing long-range motion for truncXPA compared to flXPA on $UV\lambda_{80J}$ (Figure 12d). Furthermore, the number of pause sites occupied by truncXPA on $UV\lambda_{80J}$ (37.5% of particles had only two pause sites) was dramatically less than the number of pause sites occupied by flXPA on

$UV\lambda_{80J}$ (Figure 12e). By these measures, truncXPA behavior on UV-irradiated DNA more closely resembled flXPA behavior on non-damaged DNA. Together with the biochemical assays, these data suggest that while the core DNA-binding domain of XPA is capable of recognizing damage and forming stationary complexes, its ability to transition into the paused state and form high affinity stable complexes is sharply impaired without the N- and C-termini.

3.8 XPA Changes Search Mode on the Second Time Scale

We then considered the lengths of all individual phases to obtain an estimate for the lifetimes of motile XPA particles in each mode. Although multiple comparisons resulted in statistically significant differences between some experimental groups, the majority were not significant and there appears to be no meaningful trend between substrates or XPA length (Figure 14a). Therefore, analysis was done on the combination of all flXPA data (Figure 13a). A single exponential was fit to the cumulative frequency distribution of phase lengths for each mode. The resulting mean lifetimes (τ) are as follows: 33.0 s for paused phases, 17.4 s for short-range phases, and 13.5 s for long-range phases. These data suggest that paused phases tend to last longer than phases of either diffusive mode. Moreover, within the defined limits of experimental temporal resolution, XPA changes its search state on the second time scale. We must also note the significant proportion of XPA particles that were stationary (i.e. did not move once during 300 second periods of observation, Figure 12b), which may represent a distinct population of long-lived paused XPA proteins.

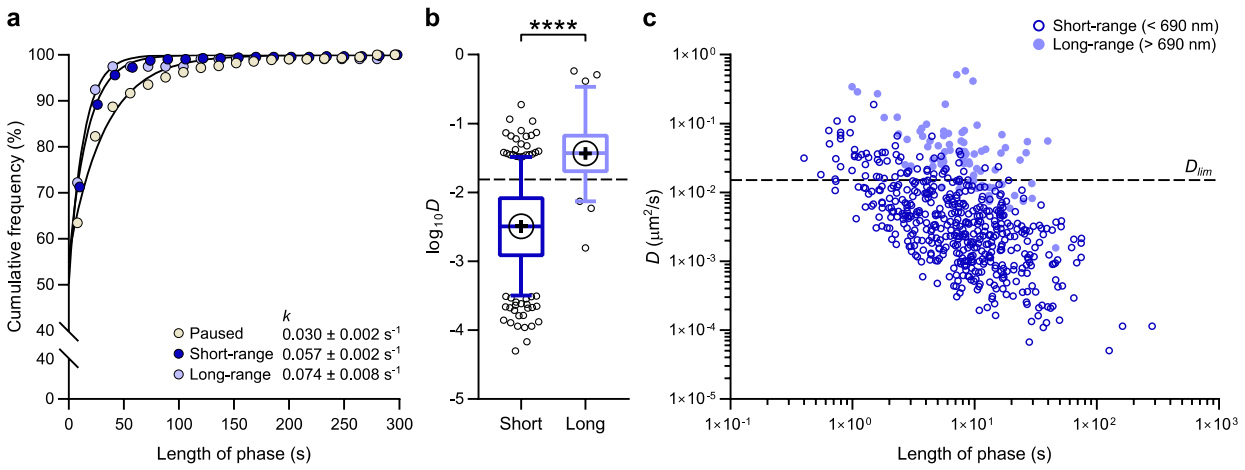

Figure 13. Short-range motion is associated with a lower diffusion coefficient.

[a] Cumulative frequency distributions of the lengths of paused (n = 638 phases), short-range (n = 665 phases), and long-range (n = 119 phases) phases for flXPA on combined λ substrates (NDλ, UVλ$_{20J}$, and UVλ$_{80J}$). A single exponential (curved line) is fit to each histogram (circles) and resulting *k* is reported as best fit value ± s.e. of the fit. Mean lifetime, $\tau = 1/k$. [b] Box and whisker plot (5-95 percentile) of the diffusion coefficient ($\log_{10}D$) calculated for short-range (n = 505 phases) and long-range (n = 79 phases) phases of Qdot-flXPA on combined λ substrates (NDλ, UVλ$_{20J}$, and UVλ$_{80J}$). +, sample mean. Dashed line, D_{lim}, theoretical limit to free diffusion for Qdot-flXPA. **** p < 0.0001 by two-tailed Student's t test. [c] Plot of diffusion coefficient (*D*) vs. length of phase (same phases shown in b). Dashed line, D_{lim}, theoretical limit to *D* for free diffusion of Qdot-flXPA.

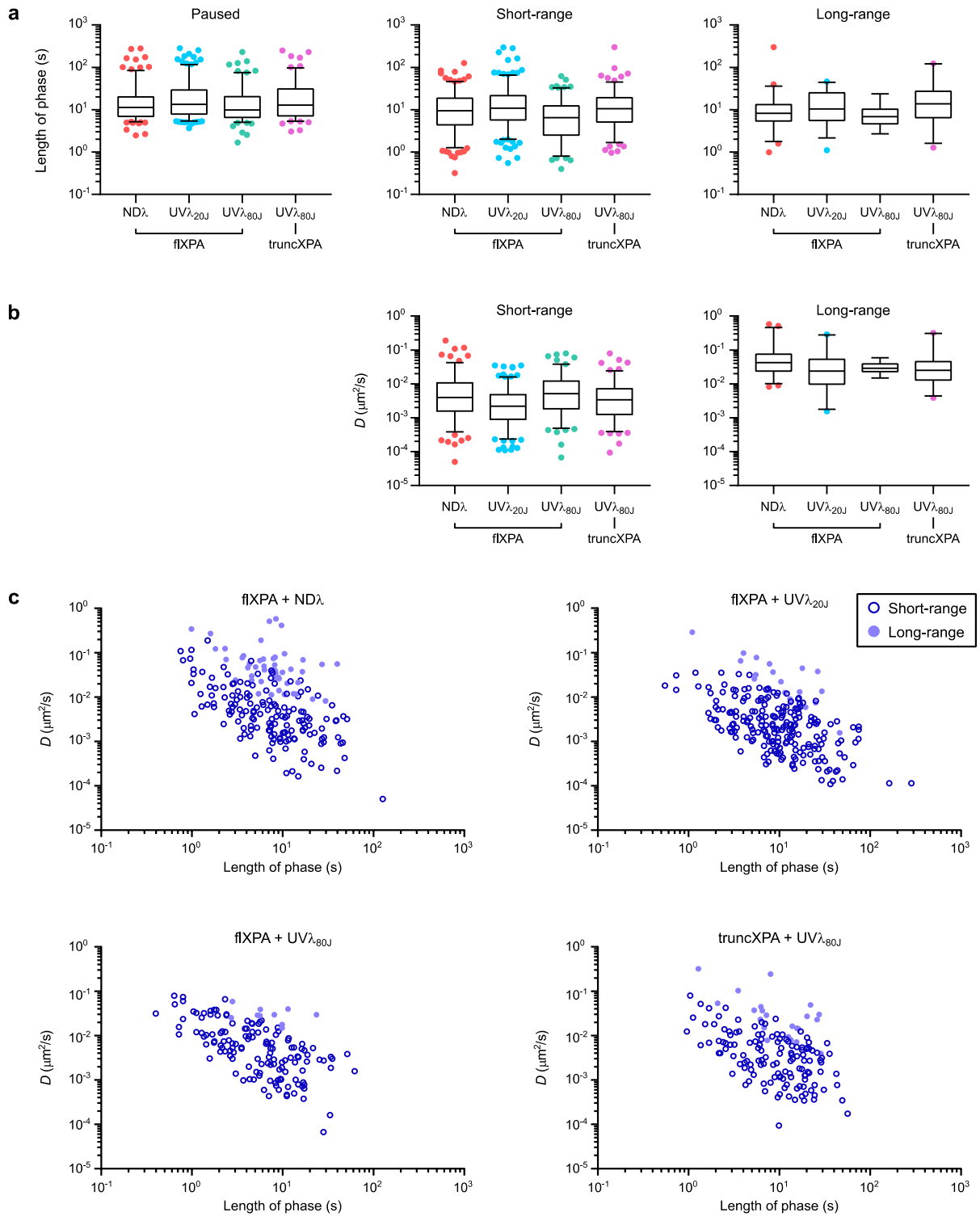

Figure 14. Comparison of phase lengths and diffusion coefficients between experiments.

[a] Box and whisker plots (5-95 percentile) of the lengths (in seconds) of all measured phases for motile XPA particles. Left, paused mode: flXPA on NDλ (n = 193 phases), flXPA on UVλ$_{20J}$ (n = 239 phases), flXPA on

UVλ_{80J} (n = 157 phases), truncXPA on UVλ_{80J} (n = 136 phases). Center, short-range mode: flXPA on NDλ (n = 214 phases), flXPA on UVλ_{20J} (n = 254 phases), flXPA on UVλ_{80J} (n = 150 phases), truncXPA on UVλ_{80J} (n = 154 phases). Right, long-range mode: flXPA on NDλ (n = 55 phases), flXPA on UVλ_{20J} (n = 27 phases), flXPA on UVλ_{80J} (n = 10 phases), truncXPA on UVλ_{80J} (n = 27 phases). All comparisons within each mode by Dunn's multiple comparisons test were not significant ($p > 0.05$) except for: paused flXPA/NDλ vs. flXPA/UVλ_{20J} ($p = 0.0025$), short-range flXPA/NDλ vs. flXPA/UVλ_{80J} ($p = 0.0005$), short-range flXPA/UVλ_{20J} vs. flXPA/UVλ_{80J} ($p < 0.0001$), and short-range flXPA/UVλ_{80J} vs. truncXPA/UVλ_{80J} ($p < 0.0001$). Data reproduced from Figure 13a, but separated to show variation between experimental conditions. [b] Box and whisker plots (5-95 percentile) of *D* of all analyzed phases. Center, short-range mode: flXPA on NDλ (n = 156 phases), flXPA on UVλ_{20J} (n = 211 phases), flXPA on UVλ_{80J} (n = 138 phases), truncXPA on UVλ_{80J} (n = 131 phases). Right, long-range mode: flXPA on NDλ (n = 49 phases), flXPA on UVλ_{20J} (n = 20 phases), flXPA on UVλ_{80J} (n = 10 phases), truncXPA on UVλ_{80J} (n = 22 phases). All comparisons within each mode by Dunn's multiple comparisons test were not significant ($p > 0.05$) except for: short-range flXPA/NDλ vs. flXPA/UVλ_{20J} ($p < 0.0001$) and short-range flXPA/UVλ_{20J} vs. flXPA/UVλ_{80J} ($p < 0.0001$). Data reproduced from Figure 13b, but separated to show variation between experimental conditions. [c] Plots of diffusion coefficient (*D*) vs. length of phase. flXPA on NDλ, n = 205 phases. flXPA on UVλ_{20J}, n = 231 phases. flXPA on UVλ_{80J}, n = 148 phases. truncXPA on UVλ_{80J}, n = 153 phases. Data reproduced from Figure 13c, but separated to show variation between experimental conditions.

3.9 Long-Range Motion is Associated with Faster Rates of Diffusion than Short-Range Motion

In order to gain insight into the rates of diffusion, we used mean squared displacement (MSD) analysis[142,167]. Because motile XPA changed its behavior so distinctly, generating MSD plots for the entire length of each kymograph was not appropriate. Instead, each diffusive phase (short- and long-range) was analyzed independently. We first combined the motions of full-length XPA particles on all substrates in order to compare diffusion between modes. The diffusion

coefficient (D) is an order of magnitude lower in the short-range mode ($D \approx 3.2 \times 10^{-3}$ μm^2/s) compared to long-range ($D \approx 3.7 \times 10^{-2}$ μm^2/s), implying that the rate of displacement is slowest when the protein is traveling shorter distances overall (Figure 13b). Interestingly, we also observed that there is a correlation between phase length and diffusion coefficient, with the fastest diffusing particles having the shortest time spent in that phase (Figure 13c). There was no meaningful difference between diffusion rates on all substrates for flXPA or truncXPA (Figure 14b); while some differences were noted between experimental groups, there was no meaningful trend, and all plots of D against phase length resulted in a similar shaped distribution (Figure 14c). This suggests that diffusion rates are inherent to the mode of diffusion itself. Together, these data indicate that while XPA's entry into a paused state is dependent upon the disordered N- and C-terminal domains, phase lengths and diffusion rates are dictated by the core DBD.

The calculated theoretical limit to the diffusion coefficient for Qdot-XPA corkscrewing along the DNA helix (D_{lim}) is 0.015 μm^2/s (see Section 2.6.5 Calculation of Theoretical Constants, Equation 9) and appears to separate the short-range from the long-range mode (Figure 13b-c). This limit was calculated with the assumption that Qdot-labeled XPA is sliding on the DNA, maintaining contact and following the helical path[153]. The fact that a significant portion of the long-range phases exceed this limit suggests that, when in this mode, XPA is undergoing an alternative mechanism of linear diffusion, namely hopping. To reconcile this, we compared the behavior of full-length XPA on UVλ_{20J} and UVλ_{80J} tightropes in buffers with different ionic strengths. Proteins that are hopping along the DNA are expected to exhibit faster diffusion in higher salt concentrations, while those that are truly sliding on DNA are expected to be relatively unaffected by changes in salt[1]. Comparing the maximum displacement of XPA, we observed a significant correlation between salt concentration and range of motion (Figure 15a). Diffusion

coefficients for XPA undergoing short-range diffusion at higher salt all fall within the range observed at 100 mM KCl (Figure 15b), thus supporting the sliding model. However, consistent with the hopping model, increased salt concentration resulted in an increase in the diffusion coefficient for XPA undergoing long-range motion.

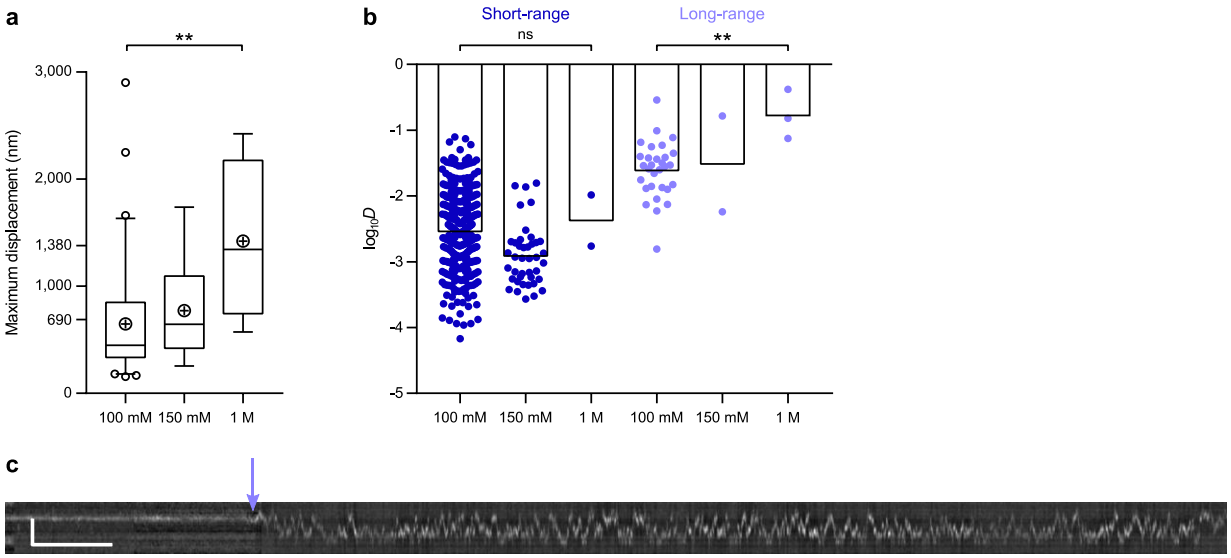

Figure 15. Effect of ionic strength on XPA diffusion.

[a] Box and whisker plot (5-95 percentile) showing maximum displacement of motile XPA particles on UVλ_{20J} and UVλ_{80J} with buffer containing 100 mM (n = 61), 150 mM (n = 11), or 1 M salt (n = 4). ⊕, sample mean. ** p = 0.0.0045 by Post test for linear trend. [b] Plot of the log transform of diffusion coefficient (D) of motile XPA on UVλ_{20J} and UVλ_{80J} with buffer containing 100 mM KCl, 150 mM NaCl, or 1 M NaCl. The 100 mM data are reproduced from Figure 13c (UVλ_{20J} and UVλ_{80J} only, subset of total), shown for comparison. Circles show individual data points, bars show means. ns, p = 0.5613; **, p = 0.0051 by Post test for linear trend. [c] Example kymograph of XPA on UVλ_{20J}. Starting buffer contains 100 mM KCl. The arrow indicates the transition from paused to long-range diffusion after the addition of 1 M NaCl. Scale bar, 2 µm (vertical) and 20 s (horizontal).

3.10 Acknowledgements

The authors wish to thank Samuel Wilson for providing purified His-Polβ and His-APE1 and Lili Liu for providing $Nick_{514}$ and purified His-UvrD. The authors also thank Muwen Kong for helpful discussions and training, Neil Kad and Hong Wang for helpful discussions, and Namrata Kumar for careful reading of the manuscript. This work was supported by the National Institutes of Health [R01ES019566 to B.V.H., R01ES028686 to B.V.H., T32GM088119 to E.C.B., and 2P30CA047904 to UPMC Hillman Cancer Center]; Deutsche Forschungsgemeinschaft [SFB-1361, TP-Carell].

4.0 Discussion

4.1 Working Model

In this study we have pursued a collection of single molecule experiments that uniquely address several questions about how XPA interacts with DNA. AFM data indicate that full-length human XPA has specificity for the helix-bending dG-C8-AAF lesion and that it binds and bends both damaged and non-damaged DNA as a monomer. Single molecule fluorescence microscopy showed that XPA is primarily stationary when bound to DNA. Of the proteins that did move, the presence of DNA damage increased pausing in a dose-dependent manner. A truncated XPA mutant consisting of just the core DNA-binding domain displayed impaired pausing compared to full-length on damaged DNA tightropes. MSD analysis revealed that the short-range mode had a significantly lower diffusion coefficient than the long-range mode, which exceeded the theoretical limit for protein sliding along the DNA contour. Our working model for XPA damage search and recognition is presented in Figure 16.

4.1.1 Stoichiometry

Although XPA has been reported to bind DNA as a homodimer on short DNA substrates[90], volumes of XPA (at concentrations between 1 and 4 μM) on a 538 bp DNA substrate measured by AFM in this study are consistent with the size of a monomer. We also observed formation of a "dimer band" at high concentrations by EMSA, but this band is indistinguishable from a complex containing one XPA bound at the lesion and one bound at the DNA end (or two otherwise distinct

monomeric binding events). Indeed, our AFM data show that a significant fraction of XPA was bound to ends of the DNA molecule in addition to specific sites. Previous work has addressed the issue of end-binding by performing protein-protein crosslinking experiments of Rad14 DBD on 15 or 37 bp AAF-adducted DNA[85]. Dimers observed in this study may be a unique property of Rad14, the truncated protein, or, as the authors state, due to unusual bending of the DNA. Furthermore, Rad14 concentrations in structural and crosslinking studies were very high (in mM amounts, compared to the μM concentrations used in the present study), as required by the protocol. As such, AFM imaging of full-length human XPA on long DNA substrates has allowed us to gain new insight into XPA binding stoichiometry. The monomeric state of free XPA observed at 40 nM by AFM and up to 80 μM by MALS provides further evidence against a dimer interface.

4.1.2 DNA Bending and Specificity

The addition of a dG-C8-AAF base modification to our AFM DNA substrate resulted in increased flexibility and a ~30° bend at that site, consistent with previous reports that an AAF modification distorts B-form DNA[165,166]. Furthermore, XPA was shown to bend both damaged and non-damaged DNA by ~60°. Extensive reports in the literature suggest that XPA binds preferentially to a distorted DNA helix[46,48,55,60,76,79,85,86,91,93-99,101-103,110], and our current work supports the hypothesis that XPA interrogates DNA by bending and testing for flexibility or pre-bent structures. The energy required to bend DNA at an AAF site is expected to be less than at a non-damaged site[32,85], and thus XPA may preferentially fold into a stable complex more readily at this and other DNA lesions.

We observed preferential XPA binding to dG-C8-AAF with a specificity factor of 660, which is in the range of reported specificities of *Taq* MutS for its substrates, a T-bulge and G:T

mismatch (1,660 and 300, respectively)[149]. Interestingly, XPA had a similar distribution of binding positions (standard deviation of the Gaussian fit was 12.3 % of the total DNA contour length) as we have previously observed for Rad4-Rad23 on a fluorescein-adducted substrate (standard deviation was 13% of the total DNA contour length)[22]. While the specificity factors determined by EMSA and AFM cannot be directly compared, in support of the AFM data, the specificity of XPA for dG-C8-AAF compared to non-damaged DNA by EMSA was ~85-fold. These levels of specificity support the "discrimination cascade" model for damage recognition in NER[58-60].

One limitation of the single molecule methods presented in this dissertation is that, by nature of the method, only relatively stable complexes can be detected. Extremely transient or weak binding events may not be recorded (e.g. appear as free unbound molecules by AFM, or never be discovered when searching for bound particles along DNA tightropes) and thus we are artificially filtering our protein population for the "best" binders. This makes it difficult to assess protein affinity or specificity for DNA targets using these methods alone. However, additional information resulting from bulk biochemical work and comparison with other well-studied proteins allows one to obtain a more complete picture of the protein-DNA interaction of interest.

4.1.3 Episodic Linear Diffusion

The role of linear diffusion in target search was explored further via the DNA tightrope assay. The episodic behavior of XPA appears to be a relatively unique property for DNA binding proteins, as few other single-particle tracking studies have yielded similar results, save for a recent report on the SA1 protein sliding on telomeric DNA tightropes[168]. It appears that XPA cycles through three distinct states on DNA: rapidly diffusing over distances greater than 2.2 kbp (690 nm), slowly diffusing and more carefully interrogating short (< 2.2 kbp) ranges of DNA, and

binding stably in a long-lived non-motile complex. Our data indicate that switching between modes is not dependent on the DNA substrate, but more likely occurs stochastically on the second time scale. These three states, combined with the AFM data, support a working model in which monomeric XPA interacts with DNA by quickly sampling large stretches of DNA by hopping, slowly interrogating and bending smaller regions as it slides along the DNA contour, and forming high-affinity complexes with sharply bent DNA (Figure 16). DNA with a high propensity for bending, such as at a dG-C8-AAF adduct[165,166] or 6-4PP[27], promotes formation of these stable complexes.

These states are reminiscent of the two-state model proposed by Slutsky and Mirny (discussed in Section 1.1)[5]. In response to the speed-stability paradox of protein-DNA target search, the authors suggest that proteins adopt two conformations: a search state with a smooth DNA-binding energy landscape, allowing for rapid search, and a recognition state with a rugged energy landscape. Comparably, linear diffusion observed by XPA fits this model. Paused and stationary particles conform to a recognition state with exceedingly rough binding energy landscapes while short- and long-range diffusing particles correspond to distinct subgroups (and thus distinct energy landscapes) within a search state.

The energy barrier to free diffusion ($\sigma = 1.6 \times k_B T$, see Section 2.6.5 Calculation of Theoretical Constants) in the short-range mode (i.e. displacement less than 690 nm or ~2.2 kbp) is essentially identical to that reported for Rad4-Rad23 undergoing constrained motion[22]. We attribute the slower rate of diffusion observed in the short-range mode to two factors. First, as XPA bends DNA to assess damage/helical distortions, this likely induces propagation of the bend along the DNA molecule, thereby increasing roughness of the energy landscape and ultimately limiting

diffusion[169]. Second, we report that XPA in this mode is translocating along the DNA by spiraling in constant contact with the helix.

Diffusion coefficients determined for XPA in the long-range mode, however, exceeded the theoretical limit of the diffusion coefficient for XPA sliding in this manner (and consequently, we are unable to calculate corresponding energy barriers to diffusion). Therefore, it is possible that XPA is hopping (i.e. rapidly undergoing micro-associations and dissociations from the DNA) in order to achieve the faster diffusion observed in the long-range mode. This model is further supported by the fact that increasing ionic strength of the buffer resulted in higher diffusion coefficients for long-range phases, presumably due to an increased distance of each "hop." The failure of increasing salt concentrations to have this effect on short-range phases again supports a sliding model, in which XPA is spiraling and maintaining contact with the DNA. In the cell, a search mechanism involving hopping would have the advantage of the protein being able to move along the DNA without being stopped by other protein-DNA complexes or nucleosomes.

4.1.4 XPA DNA-Binding Domain

We next hypothesized that the intrinsically disordered N- and C-termini of XPA play a role in DNA bending and pausing. Here we show that truncated XPA, consisting of just residues 98-239 of the core DNA-binding domain, maintains its ability to adopt a recognition state on damaged DNA. No differences were observed in phase lengths or diffusion coefficients between the truncated and full-length proteins. The number of stationary particles (i.e. long-lived non-motile complexes) and the lengths of each paused phase were statistically non-distinct between truncXPA and flXPA on UV-damaged DNA tightropes, indicating that both proteins are able to form stable complexes on DNA and that once this complex is formed, the N- and C-terminal arms are not

involved in switching modes out of the paused state. However, motile particles of truncated XPA were significantly impaired in their ability to pause on damaged DNA, as demonstrated by a decreased number of pause sites as well as decreased total time spent in the paused mode (as a result of a lower frequency of paused phases). Therefore, the N- and C-terminal arms likely do play a role in finding the lesion and may serve as metaphorical brakes, allowing XPA to slow down and test the DNA for a lesion. This is further supported by EMSA results showing that truncXPA binds non-damaged DNA with similar affinity as flXPA, but has approximately half of the specificity for an AAF adduct. These data are consistent with a model in which the core DNA-binding domain is able to form stable complexes, but the disordered arms play a role in embracing the DNA and interrogating for helix-distorting damage, inducing a conformational change to stabilize pausing.

An important study by Chazin and colleagues concluded that residues 98-239 of XPA were sufficient to achieve wild type DNA-binding[77]. They demonstrated that truncated XPA, containing just these residues, bound dsDNA, ssDNA, and a Y-shaped ss/dsDNA junction with similar affinity as full-length XPA. It is important to note that DNA duplexes representing initial NER substrates/lesions were not tested. As previously reported, we observed similar affinity between truncXPA and flXPA for non-damaged dsDNA. However, we show that residues 98-239 were not sufficient to achieve wild-type binding to an AAF-adduct. This distinction suggests that perhaps the N- and C- termini are not required for XPA to bind the Y-shaped substrate[77], but they do play a role in recognizing and stopping at a lesion. This is consistent with a model in which the disordered arms of the protein are involved in the early steps of damage recognition. Then, as NER proceeds and XPA must bind to intermediate DNA structures and act as a scaffold protein, the arms change function and interact with other proteins rather than DNA.

Along these lines, given the many reported interactions between XPA and other NER proteins[54,58], we cannot ignore the possibility that XPA behaves differently in the presence of other repair partners. As such, we present our model as a starting point, illustrating XPA DNA binding behavior and diffusive properties in isolation, with further studies required to investigate the role and impact of its interaction partners.

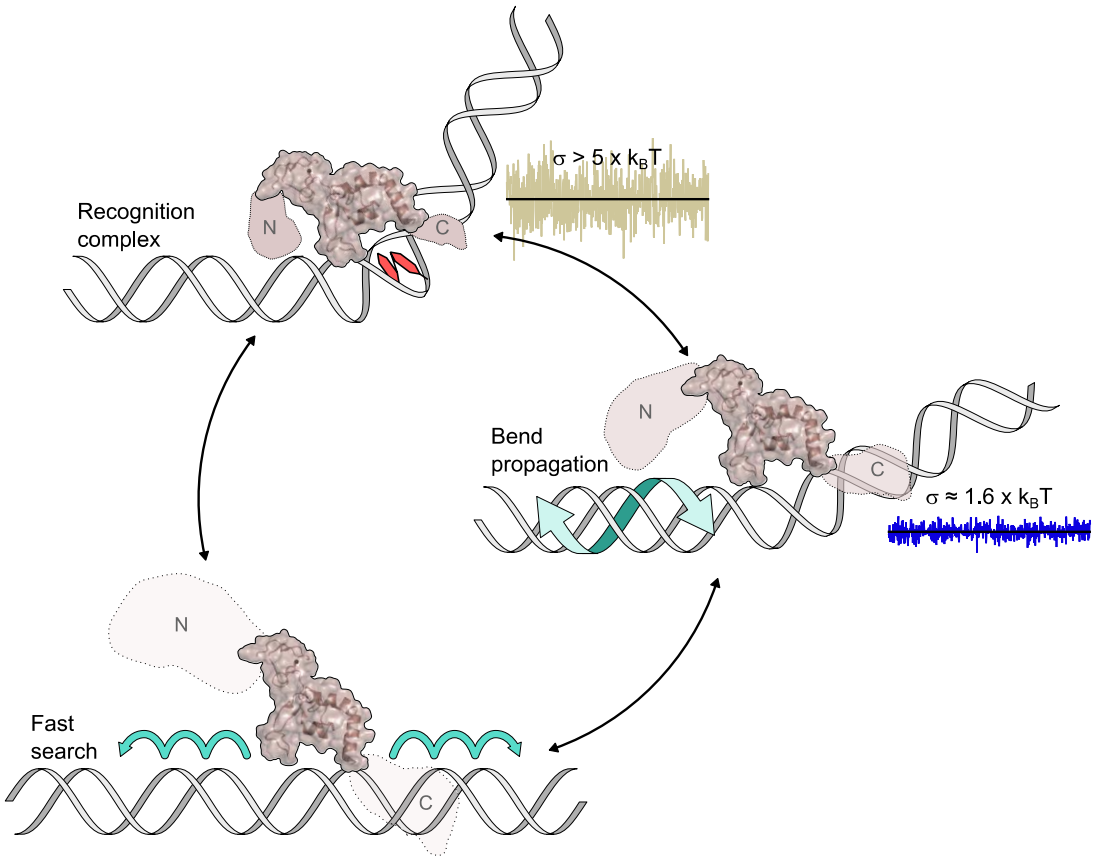

Figure 16. Working model of XPA linear diffusion on DNA.

<u>Fast search</u> (long-range linear diffusion): monomeric XPA translocates rapidly by hopping (micro-associations/dissociations) along the DNA backbone. The protein's disorder and lack of contacts with DNA permit rapid diffusion (D ≈ 0.04 μm^2/s). <u>Bend propagation</u> (short-range linear diffusion): XPA follows the path of the major or minor groove, spiraling along DNA (D ≈ 0.003 μm^2/s). The N- and C-termini of the protein may be partially folded, making more contacts with the DNA. The protein bends DNA, testing for flexibility/aberrations, and propagation of the bend further slows diffusion. <u>Recognition complex</u> (paused and stationary particles): the disordered ends of the protein fold into a stable complex, bending the DNA ~60°. The NMR structure of the XPA DNA-binding domain (PDB 1XPA) was used in combination with an estimation for the N- and C-termini (disorder represented by size/transparency) as an approximate model to display the complete XPA molecule.

4.2 Outlook

The single molecule studies of XPA on DNA presented here resolve several important questions about its damage recognition behavior (as discussed in Section 4.1). However, in the everlasting cycle of scientific research, these studies have perhaps opened more questions than they have answered.

4.2.1 XPA Specificity

As laid out in Appendix B, Table 1, and discussed in Section 1.3.2 (XPA-DNA Interactions), there have been numerous studies investigating XPA specificity for a wide range of DNA substrates. However, none have done a side-by-side comparison with a truly helix-bending lesion in dsDNA and forked NER-intermediate mimics. There is much debate and conflicting conclusions in the literature regarding XPA's role in damage recognition. The "early XPA" supporters tend to cite literature showing XPA has specificity for DNA lesions while "late XPA" supporters give more weight to literature examining forked substrates, and it is significantly difficult to compare results between papers. The common theme with all these studies is that XPA binds flexible DNA preferentially over dsDNA. If these studies are going to be used to support one model vs. another, then perhaps one way to resolve some of the conflicting reports would be to do an extensive and comprehensive comparison study of full-length XPA binding to a wide range of DNA substrates, including multiple bulky adducts, crosslinks, and ss/dsDNA junctions.

Sequential assembly of NER proteins has also been studied in cells by immunofluorescence (IF). Vermeulen and colleagues showed that XPA is required to recruit ERCC1 to UV-damaged foci, but not XPC, TFIIH, or XPG[47]. Such studies are more technically challenging to study

recruitment to chemical base damage like dG-C8-AAF. Furthermore, although the evidence demonstrates that XPA may not be required for recruitment of all downstream proteins, it does not negate the possibility that XPA may enhance damage recognition efficiency or that its interaction with DNA lesions is important for damage verification. Single molecule tracking of XPA and other recognition proteins (UV-DDB or XPC) in living cells would provide insight into complex formation in real time.

4.2.2 Role of DNA Bending in XPA Damage Recognition

Results of AFM experiments with XPA binding to dsDNA with or without an AAF adduct indicate that DNA bending plays an important role in recognition and target search. Currently, we have only examined DNA bending using our AFM platform, which provides a static snapshot of molecular conformations. We would like to resolve four additional questions stemming from these results.

First, does XPA bind preferentially to DNA already in a sharply (60°) bent state (i.e. conformational capture) or does it induce conformational change of the DNA upon binding? AFM imaging of free DNA molecules with or without the AAF adduct indicate that the majority of molecules do not adopt a bend angle greater than 34° or 10°, respectively. However, as can be seen in Figure 11, a small number of unbound DNA molecules had a bend greater than 60°. Is XPA finding these pre-bent structures or does the DNA bend after binding? This is a subtle distinction, but the use of time lapse or high-speed AFM, which allows for greater temporal resolution, may allow one to observe protein-induced DNA bending, should it occur. Because XPA has an established propensity to bind DNA lesions which are both bent and flexible, I hypothesize that XPA does bind preferentially to bent structures, but with an additional component of induced fit,

resulting in a sharply bent structure. Along these lines, a parallel idea for damage recognition by UV-DDB, namely conformational proofreading, has been proposed[21]. The FQH hairpin of DDB2 probes the major groove of the DNA molecule for damage and encounter with an appropriate lesion induces conformational changes in both protein (formation of an α paddle) and DNA (base flipping)[21,170].

Second, does XPA have any specificity for other non-damage associated bends in DNA? While numerous studies have examined XPA binding to a range of artificially distorted DNA substrates (see Section 1.3.2 and Appendix B), none have looked at bends caused by A-tract:GC junctions in dsDNA. We could test this first by EMSA, but perhaps more effectively by AFM. By measuring the XPA binding position on a 538 bp dsDNA fragment with this specific sequence at 30% from one end would reveal any preferential binding. Affinity for sequence-induced DNA bending may also help to understand XPA pausing on non-damaged DNA tightropes (see below, Section 4.2.3 XPA Specificity on Non-Damaged DNA).

Third, how does DNA tension and range of bending impact XPA diffusion on DNA tightropes? Under the current experimental set-up, DNA tightropes are suspended at 90% the total contour length of the DNA. These conditions should allow for localized DNA bending and have been demonstrably acceptable for the study of a number of DNA-binding proteins[21,23], including those that bend DNA[22]. However, it is possible that reducing DNA tension even further may allow DNA bending more readily, and thus enhance XPA binding. To test this using our DNA tightrope platform, we would string up DNA under slower flow to allow for gentler elongation of DNA molecules. Investigating XPA diffusion on DNA tightropes at 80% or 70% contour length may reveal increased pausing. One limitation of this kind of experiment is that, if DNA tightropes are given too much slack, they will not remain within one plane (XY or Z) and will constantly move

in and out of focus when imaging. An alternative approach using optical tweezers may avoid this problem. In this kind of experiment, optical tweezers would first trap a DNA molecule at both ends, and then XPA would bind the DNA, and finally the tweezers would be manipulated to bring the ends of the molecule closer or further apart. By combining optical tweezers with fluorescence microscopy, use of the C-Trap™ (Lumicks) with labeled XPA would allow for direct visualization of this in real time. If our hypothesis that XPA must bend the DNA in order to form a stable complex is correct, we expect XPA to dissociate from DNA molecules as the ends are stretched further apart.

Fourth, can the truncated XPA mutant bend DNA? Our single molecule tightrope experiments showed that, without its disordered N- and C-termini, XPA exhibits reduced pausing on damaged DNA. When it does pause, it stays paused for the same length of time as full-length XPA. If DNA bending is essential for specificity and pausing by XPA, then I predict that the truncated mutant will show impaired DNA bending. We can use AFM to measure DNA bend angles at sites of bound truncXPA on both the non-damaged and AAF-adducted substrates. The full-length XPA protein bent DNA ~60° at both non-specific and specific sites. If the N- and C-terminal arms are necessary for DNA bending, we should see a smaller bend angle or perhaps a wider distribution of angles.

4.2.3 XPA Specificity on Non-Damaged DNA

Analysis of the number of pause sites occupied by XPA on various DNA tightropes indicated that XPA tends to pause at UV-induced lesions. In addition to this, we also observed XPA pausing on NDλ tightropes. What is the nature of the pause sites on NDλ?

To answer this, chromatin immunoprecipitation (ChIP) of different DNA samples may show XPA binding preferentially to certain genomic sequences. Specifically, as tightrope assays were performed using λ DNA, it is of most interest to determine if XPA is enriched at any particular regions of this substrate. Purified XPA (full-length, with a His tag) would be incubated with and crosslinked to λ DNA, the DNA would be fragmented, and XPA-DNA complexes would be immunoprecipitated using beads with anti-His antibodies attached. Adapter ligation may be used to build a library of DNA fragments for sequencing. The benefit of this assay is that it makes no assumptions about what XPA is seeing when it binds.

However, because we do have additional data to inform our hypotheses, we can also begin to make some assumptions. As discussed above (Section 4.2.2 Role of DNA Bending in XPA Damage Recognition), XPA may be pausing at sequence-induced DNA bends. Alternatively, XPA may be pausing at spontaneous damage (most likely nicks in the DNA backbone or oxidized bases). ChIP sequencing results may also indicate other preferred sequences. Biochemical (EMSA or otherwise) assays in combination with AFM may be used to determine XPA specificity for such sequences and/or lesions to support ChIP sequencing results.

Finally, these alternative specific binding sites for XPA may be tested using the DNA tightrope assay. Using NDλ DNA tightropes, DNA sequences, in close proximity to the predicted pause site, may be labeled with a biotinylated oligonucleotide, which can form a DNA triplex and subsequently be labeled with a streptavidin-coated Qdot. If the predicted sequences are true XPA pause sites, then colocalization should be observed between the labeled DNA and XPA, labeled orthogonally with a different color Qdot. Alternatively, arrays of plasmids with short sequences representing the predicted pause site and a proximal biotinylated base can be prepared as described (Appendix F). Colocalization would be observed in the same way. In the proposed experiment,

XPA may have limited range of linear diffusion if the DNA label serves as a roadblock to sliding (although XPA undergoing hopping would be expected to bypass the triplex DNA). Therefore, controls must be performed with DNA labeled at a "neutral" site to test if XPA will pause at the label, regardless of DNA sequence.

4.2.4 XPA Interaction with Other NER Proteins

As discussed in Section 1.3.3 (XPA-Protein Interactions), XPA is an important scaffold protein for NER. The studies presented in this dissertation, using purified XPA and DNA in isolation, were necessary to gain a fundamental understanding for DNA damage search by XPA. They also provide the groundwork for subsequent studies with additional proteins. Ultimately, we would like to fully reconstitute mammalian NER at the single molecule level.

To this end, the next steps, with respect to XPA, would be to examine how XPA interaction with damaged DNA is altered in the presence of partner proteins. Specifically, UV-DDB[52] and RPA[55] have both been reported to enhance binding affinity of XPA for damaged DNA. Therefore, I predict that addition of these proteins to enhance damage search by XPA. Using orthogonally-labeled proteins on our DNA tightrope platform, one could look for colocalization of XPA with UV-DDB or RPA on DNA. Furthermore, any impact on diffusion, such as increased frequency of stationary particles or pausing within the motile fraction, could be observed as a synergistic target search. AFM could also be used to observe formation of ternary complexes on DNA.

4.2.5 XPA Damage Search in Chromatin

One last set of experiments (though certainly not the requisite end of this story) would investigate the effects of chromatin structure on damage search by XPA. The use of long DNA substrates in our tightrope assay is arguably a more biologically relevant substrate compared to the short oligonucleotides often used in biochemical assays. However, DNA in human nuclei is wrapped in nucleosomes and condensed into chromatin. Histone proteins may thus act as roadblocks or physical barriers to XPA sliding along DNA. Finkelstein and colleagues recently used a single molecule tracking platform (DNA curtains[171-173]) to show that the yeast mismatch repair proteins Msh2-Msh3 primarily search DNA for damage via a hopping mechanism, and thus are able to bypass nucleosomes during target search[174]. Conversely, they found that Msh2-Msh6 performs target search via sliding, and cannot hop over nucleosomes[174].

In this manner, by adding nucleosomes to our DNA tightrope platform, we can investigate XPA diffusion in the context of chromatin. The ability of XPA to perform facilitated diffusion by alternating between sliding and hopping may allow for the protein to bypass nucleosomes when undergoing long-range diffusion, but not short-range. Thus, we would expect to see particles undergoing long-range diffusion to continue past such roadblocks. The range of displacement for particles undergoing short-range diffusion would be limited by the spacing between nucleosomes.

4.3 Concluding Remarks

In this dissertation, we present a thorough single molecule analysis of XPA DNA damage recognition. Taken together, our data is consistent with a three-state model for DNA target search

(Figure 16). The “search state” of XPA can be broken down into two modes. Fast search corresponds to a conformation which allows for rapid sampling over long (> 690 nm) distances of DNA, which involves DNA hopping. Slow search corresponds to a conformation with slower linear diffusion via sliding over short distances. In this mode, we predict XPA to bend DNA and test for flexibility. The “recognition state” of XPA occurs when the DNA helix is bent by ~60°. The intrinsically disordered N- and C-termini of the protein embrace the DNA and facilitate pausing and formation of the recognition state.

Appendix A Abbreviations

λ – Enterobacteria phage lambda

6-4PP – (6-4) pyrimidine-pyrimidone photoproduct

AF – 2-aminofluorene

AAF – 2-acetylaminofluorene

AAN – N^2-acetylnaphthyl

AFM – Atomic force microscopy

AP – 1-aminopyrene

APE1 – Apurinic/apyrimidinic endonuclease 1

B[a]P – Benzo[a]pyrene

BER – Base excision repair

bp – Base pair

ChIP – Chromatin immunoprecipitation

CPD – Cyclobutane pyrimidine dimer

CS – Cockayne syndrome

CSA – Cockayne syndrome protein A

CSB – Cockayne syndrome protein B

C-terminus – Carboxy-terminus of a protein

D – Diffusion coefficient

D_{lim} – Theoretical limit to the diffusion coefficient

DBD – DNA-binding domain

DNA – Deoxyribonucleic acid

dG-C8-AAF – N-(2’-deoxyguanosin-8-yl)-2-acetylaminofluorene

dsDNA – Double-stranded DNA

ELISA – Enzyme-linked immunosorbent assay

EMSA – Electrophoretic mobility shift assay

FITC – Fluorescein isothiocyanate

FL – Fluorescein

FRET – Fluorescence resonance energy transfer

GG NER – Global genome nucleotide excision repair

HMGB1 – High mobility group protein B1

IDT – Immobilized DNA template

IF – Immunofluorescence

IPTG – Isopropyl-β-D-thiogalactoside

k_BT – Thermal energy term, product of Boltzmann constant and temperature

K_D – Equilibrium dissociation constant

MALS – Multiple angle light scattering

MM – DNA mismatch

MMC – Mitomycin C

MSD – Mean squared displacement

ND – Non-damaged

NER – Nucleotide excision repair

nt – Nucleotide

N-terminus – Amino-terminus of a protein

PCNA – Proliferating cell nuclear antigen

Polβ – DNA polymerase β

Qdot – Quantum dot

RFC – Replication factor C

RPA – Replication protein A

SEC – Size exclusion chromatography

SPR – Surface plasmon resonance

ssDNA – Single-stranded DNA

TC NER – Transcription-coupled nucleotide excision repair

TIRF – Total internal reflection fluorescence

TFIIH – Transcription factor IIH

TTD – Trichothiodystrophy

TTDA – Trichothiodystrophy group A protein, also called p8

UV – Ultraviolet

UV-DDB – UV-damaged DNA binding protein

XAB1 – XPA binding protein 1

XAB2 – XPA binding protein 2

XL – Crosslink

XP – Xeroderma pigmentosum

XP-A – Xeroderma pigmentosum, complementation group A

XPA – DNA repair protein complementing XP-A cells

ZF – Zinc finger

Appendix B XPA-DNA Interactions

Table 1. Published XPA-DNA interactions.

Unless otherwise noted: studies were performed using full-length wild-type human protein (recombinant, purified), fold specificities are reported as overall binding between substrates, and DNA substrates were prepared using short oligonucleotides (less than 60 bp or nt).

Paper	Substrate	K_D	Specificity	Method
Robins et al, 1991.[93]	UV-irradiated dsDNA (9 kJ/m^2) UV-irradiated dsDNA (0 - 9 kJ/m^2) Non-damaged ssDNA Non-damaged dsDNA		~1,000-fold specificity[a] for UV damage over non-damaged dsDNA. Fluence-dependent affinity for UV damage. Lower affinity for ssDNA than for dsDNA.	Filter binding[b,c]
Jones and Wood, 1993.[94]	UV-irradiated dsDNA (6 kJ/m^2) UV-irradiated dsDNA (0-6 kJ/m^2), treated with CPD photolyase Cisplatin-treated dsDNA Psoralen-treated dsDNA Non-damaged dsDNA Non-damaged ssDNA, circular Non-damaged dsDNA, circular	333 nM 1.67 μM 	~300-fold specificity[a] for 6-4PP over dsDNA. ~4-fold specificity for circular ssDNA over circular dsDNA. Higher affinity for cisplatin than for non-damaged dsDNA. No specificity for CPD or psoralen adducts.	EMSA[d]

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Asahina et al, 1994.[79]	UV-irradiated dsDNA (8 kJ/m^2) Cisplatin-treated dsDNA OsO_4-treated dsDNA Non-damaged ssDNA Non-damaged dsDNA		Specificity for UV-treated DNA over non-damaged. Specificity for dsDNA over ssDNA. Higher affinity for cisplatin than OsO_4-treated DNA; specificity for both over non-damaged.	Filter binding[e]
Li et al, 1995.[55]	UV-irradiated dsDNA (600 J/m^2) Non-damaged dsDNA		Specificity for UV-treated DNA over non-damaged.	IDT[f]
Kuraoka et al, 1996.[76]	UV-irradiated dsDNA (8 kJ/m^2) Cisplatin-treated dsDNA Non-damaged dsDNA		Specificity for UV-treated and cisplatin-treated DNA over non-damaged.	Filter binding[g]
Nocentini et al, 1997.[48]	UV-irradiated dsDNA (1 kJ/m^2) Non-damaged dsDNA		Specificity for UV-treated DNA over non-damaged.	Filter binding

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Buschta-Hedayat et al, 1999.[98]	AAF-adducted dsDNA (-)-cis-B[a]P-adducted dsDNA (-)-trans-B[a]P-adducted dsDNA dsDNA with 3 nt MM dsDNA with 1 nt MM C4' pivaloyl-adducted dsDNA C4' pivaloyl-adducted dsDNA, adduct in 3 nt MM 3-nitropyrrole-modified dsDNA 5-nitroindole-modified dsDNA Non-damaged ssDNA Non-damaged dsDNA		Specificity for AAF and B[a]P over non-damaged dsDNA. Specificity for C4' pivaloyl adduct within bubble, but not in dsDNA. Higher affinity for 3 nt mismatch than 1 nt mismatch, specificity for both over dsDNA. Lower affinity for ssDNA than for dsDNA. Higher affinity for 5-nitroindoles than for 3-nitropyrroles, specificity for both over non-modified dsDNA.	EMSA
Wakasugi et al, 1999.[46]	6-4PP-modified dsDNA Non-damaged dsDNA	6 nM 420 nM	~70-fold specificity[a] for UV-treated DNA over non-damaged.	EMSA

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Wang et al, 2000.[97]	6-4PP-modified dsDNA	21 nM	~4.5-fold specificity for ssDNA over dsDNA. ~3-fold specificity for 6-4PP. ~1.3-fold specificity for CPD.	SPR
	CPD-modified dsDNA	46 nM		
	Non-damaged ssDNA	13 nM		
	Non-damaged dsDNA	58 nM		
Mustra et al, 2001.[110]	dsDNA with MMC interstrand XL		~2-3 fold specificity for MMS crosslink over non-damaged dsDNA.	EMSA
	Non-damaged dsDNA			
Hey et al, 2001.[96]	Cisplatin-adducted dsDNA, 3' FL	415 nM	~3-fold specificity for cisplatin, ssDNA loop, mismatched bubble, and ssDNA with mixed bases over non-damaged dsDNA. ~1.5-fold specificity for pyrimidine-rich ssDNA. No specificity/worse binding to purine-rich ssDNA compared to dsDNA.	Anisotropy
	dsDNA with 6 nt MM, 3' FL	380 nM		
	dsDNA with 3 nt insert on one strand, 3' FL	350 nM		
	Non-damaged ssDNA, mixed bases	355 nM		
	Non-damaged ssDNA, AG-rich	> 3 μM		
	Non-damaged ssDNA, TC-rich	786 nM		
	Non-damaged dsDNA, 3' FL	1.15 μM		

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Missura et al, 2001.[102]	Cisplatin-adducted dsDNA Dinuclear cisplatin analogue-adducted dsDNA (Pt-Pt) dsDNA with 3 nt MM dsDNA with 3 nt insert on one strand dsDNA with 3 bp insert on one strand Y shaped DNA 3-way dsDNA junction 4-way dsDNA junction 3-nitropyrrole-modified dsDNA 5-nitroindole-modified dsDNA Non-damaged ssDNA Non-damaged dsDNA		Specificity for cisplatin, but not for the dinuclear analogue, over non-damaged DNA. Specificity for non-hybridized substrates as follows: dsDNA insert > ssDNA insert > mismatch bubble > dsDNA. Specificity for forked substrates as follows: 4-way dsDNA > 3-way dsDNA > Y > non-damaged dsDNA. Authors note "extraordinary affinity" of XPA for 4-way and 3-way dsDNA junctions. Higher affinity for 5-nitroindoles than for 3-nitropyrroles, specificity for both over non-modified dsDNA. No affinity for ssDNA.	EMSA
Iakoucheva et al, 2002.[95]	dsDNA with 4 nt MM, 5' FL	158 nM	~5-fold lower affinity for mismatch compared to dsDNA.	Stop flow[h]
	Non-damaged dsDNA, 5' FL	28.9 nM		
	Non-damaged dsDNA, 5' FL	24.4 nM		Equilibrium fluorescence[h]

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Reardon and Sancar, 2003.[60]	6-4PP-modified dsDNA CPD-modified dsDNA Non-damaged dsDNA	150 nM 210 nM 220 nM	~1.5-fold specificity for 6-4PP over non-damaged dsDNA. No specificity for CPD. Note: authors report similar fold specificity for RPA and XPC on same substrates.	EMSA
Liu et al, 2005.[90]	AAF-adducted dsDNA, 5' FL	714 nM (K_{D1}), 55 nM (K_{D2})	Note: authors report positive cooperativity (Hill = 1.9)	Anisotropy
	AAF-adducted dsDNA Non-damaged dsDNA	200 nM	Higher affinity for dG-C8-AAF than for non-damaged dsDNA.	EMSA
Brabec et al, 2006.[101]	Cisplatin-adducted dsDNA,1,3-GTG Cisplatin-adducted dsDNA, 1,2-GG Non-damaged dsDNA		4-5-fold specificity for 1,2-GG adducts, when flanked T or A bases, over non-damaged dsDNA. ~2-fold specificity for 1,2-GG when flanked by C's. Less than 2-fold specificity for 1,3-GTG adducts.	EMSA
Camenisch et al, 2006.[103]	Cisplatin-adducted dsDNA Y shaped DNA 3-way dsDNA junction 4-way dsDNA junction Non-damaged dsDNA		Specificity for cisplatin over non-damaged dsDNA. Specificity for forked substrates as follows: 4-way dsDNA > 3-way dsDNA > Y > non-damaged dsDNA.	EMSA

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Yang et al, 2006.[100]	Y-shaped DNA	49 nM		Anisotropy
	5' overhang		Similar affinity for 3' overhang and 5' overhang. Similar affinity for all mismatch bubbles, with or without lesion. Higher affinity for bubbles with 8 or more mismatched bases. Specificity for G[8,5-Me]T crosslink over non-damaged DNA. No specificity for intrastrand crosslinks formed by carbon tethers. No affinity for non-damaged ssDNA or dsDNA.	EMSA
	3' overhang			
	Y-shaped DNA			
	dsDNA with 6 nt MM			
	AF-adducted DNA, lesion in 6 nt MM			
	AAF-adducted DNA, lesion in 6 nt MM			
	AP-adducted DNA, lesion in 6 nt MM			
	6-4PP-modified DNA, lesion in 6 nt MM			
	dsDNA with 3, 4, 5, 8, 10, or 12 nt MM			
	G[8,5-Me]T XLed dsDNA			
	dsDNA with two-carbon tether XL at GG			
	dsDNA with three-carbon tether XL at GG			
	dsDNA with four-carbon tether XL at GG			
	Non-damaged ssDNA			
	Non-damaged dsDNA			

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Mustra et al, 2007.[91]	dsDNA with MMC interstrand XL Non-damaged dsDNA		~2-fold specificity for MMC XL over non-damaged DNA.	EMSA
Krasikova et al, 2008.[116]	FL-dUMP-adducted dsDNA		Specificity for Flu-dUMP over non-damaged dsDNA	EMSA
Brown et al, 2010.[99]	AAF-adducted dsDNA Thymine glycol-modified dsDNA Non-damaged dsDNA	44 nM 48 nM	Similar affinity for AAF and thymine glycol. Specificity for both over non-damaged dsDNA.	EMSA
Sugitani et al, 2014.[77]	Y-shaped DNA, 5' FL (label at dsDNA end) Non-damaged ssDNA, 5' FL Non-damaged dsDNA, 5' FL	290 nM 1.5 μM 1.7 μM	~6-fold specificity for Y-shaped DNA over ssDNA and dsDNA.	Anisotropy

Table 1 continued

Paper	Substrate	K_D	Specificity	Method
Koch et al, 2015.[85]	AAF-adducted dsDNA FITC-adducted dsDNA Cisplatin-adducted dsDNA		Higher affinity for dG-C8-AAF and FITC than for cisplatin.	EMSA
	AAF-adducted dsDNA FITC-adducted dsDNA Cisplatin-adducted dsDNA 6-4PP-modified dsDNA CPD-modified dsDNA	135 nM	No specificity for CPD or 6-4PP.	EMSA[i]
Ebert et al, 2017.[86]	AF-adducted dsDNA AAF-adducted dsDNA AAN-adducted dsDNA Non-damaged dsDNA		Higher affinity for dG-C8-AAF than for dG-N^2-AAN or dG-C8-AF. Higher affinity for all lesions compared to non-damaged dsDNA.	EMSA

Abbreviations, see Appendix A.

[a] Specificity calculated to account for nonspecific bases in damaged substrate[94]

[b] XPA fractionated from calf thymus

[c] 779 bp and 2961 bp (mixed) DNA

[d] Linear DNA substrates, 258 bp; circular substrates, M13 DNA

[e] 7250 bp DNA

[f] 622 bp and 485 bp (mixed) DNA

[g] 2686 bp DNA

[h] *Xenopus laevis* XPA
[i] *Saccharomyces cerevisiae* Rad14 (residues 10-end)

Appendix C XPA-Protein Interactions

Table 2. Published XPA-protein interactions.

Protein/ complex	Step in NER	XPA residues involved	Notes	Methods, Refs.
UV-DDB	Damage recognition	**DDB2**: 185-226	Interact both on and off DNA. Interaction enhances damage binding by both proteins. XPA R207G mutant abolishes interaction.	EMSA (supershift)[52] DNase protection[52] Pull-down[113] Immunofluorescence[113]
XPC	Damage recognition	Unknown	Proteins do not interact together on DNA.	Pull-down[114,115] SPR[114] Protease protection[115] EMSA (supershift)[116]
TFIIH	Damage verification	**XPB**: N- and C-termini of DBD (~160 and 227-239) **XPD**: N-terminus (~66-170) **p52**: C-terminus (~239-262) **p8/TTDA**: C-terminus	Interact on DNA, and weakly off of DNA. XPA forms bridge between XPB, XPD, and p52, likely extending to p8/TTDA. Interaction not observed by EMSA supershift. PDB 5OF4.	Pull-down[48,118,119] EMSA (supershift)[118] Cryo-electron microscopy[87] Crosslinking[87]

Table 2 continued

Protein/ complex	Step in NER	XPA residues involved	Notes	Methods, Refs.
RPA	Damage verification/ incision	**RPA32**: 29-46 **RPA70**: 102-176	Interact both on and off DNA. XPA interacts directly with RPA32 and RPA70, but not RPA14. K_D for RPA binding to XPA = 19 nM (ref. [122]).	Pull-down[56,57,120-122] Affinity chromatography[51,55,120] ELISA[123] Yeast two-hybrid[55,57] SPR[97,122] NMR titration[50,51,83,124] Fluorescence anisotropy[96]
ERCC1	Incision	67-80	Interact both on and off of DNA. Involved in recruitment of XPF/ERCC1 to DNA. K_D for ERCC1 binding to XPA = 250 nM (ref. [122]). PDB 2JNW.	Yeast two-hybrid[49,126] Pull-down[49,122,126,175,176] NMR[176] Affinity chromatography[177,178] SPR[122] Sedimentation equilibrium[178] X-ray crystallography[178]
PCNA	Gap filling	161-170	Connection between NER and replisome.	Immunofluorescence[53] FRET[53]
ATR	n/a	98-219	ATR phosphorylates XPA S196. Enhances stability and nuclear import of XPA.	Immunofluorescence[179,180] Pull-down[180] Protein footprinting[180]
HMGB1	n/a	Unknown	Interactions may be indirect or dependent on other factors present in cell extracts.	Pull-down[181] Chromatin immunoprecipitation[182]

Table 2 continued

Protein/ complex	Step in NER	XPA residues involved	Notes	Methods, Refs.
PARP1	n/a	213-237	PAR-dependent interaction.	Pull-down[183]
TFIIE	n/a	Unknown	XPA interacts with p34 subunit, but not p56-p34 complex. TFIIE does not appear to be involved in NER.	Pull-down[118]
XAB1	n/a	N-terminus	XPA-binding protein 1, GTPase.	Yeast two-hybrid[184]
XAB2	n/a	Unknown	XPA-binding protein 2, suggested role in transcription and TC NER.	Yeast two-hybrid[185] Pull-down[185]

Appendix D Studying Protein-DNA Interactions Using Atomic Force Microscopy

Review of AFM studies on protein-DNA interactions, originally published in *Seminars in Cell and Developmental Biology*. Ref. [129]: Beckwitt, E. C., Kong, M. & Van Houten, B. Studying protein-DNA interactions using atomic force microscopy. *Semin Cell Dev Biol* **73**, 220-230, doi:10.1016/j.semcdb.2017.06.028 (2018).

Seminars in Cell & Developmental Biology 73 (2018) 220–230

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Studying protein-DNA interactions using atomic force microscopy

Emily C. Beckwitt[a,b], Muwen Kong[a,b], Bennett Van Houten[a,b,c,*]

[a] Program in Molecular Biophysics and Structural Biology, University of Pittsburgh, Pittsburgh, PA 15261, USA
[b] The University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, PA 15213, USA
[c] Department of Pharmacology and Chemical Biology, University of Pittsburgh, Pittsburgh, PA 15261, USA

ARTICLE INFO

Article history:
Received 3 May 2017
Received in revised form 27 June 2017
Accepted 29 June 2017
Available online 30 June 2017

Keywords:
Atomic force microscopy
Protein-DNA interaction
Binding specificity
DNA bending
Stoichiometry

ABSTRACT

Atomic force microscopy (AFM) has made significant contributions to the study of protein-DNA interactions by making it possible to topographically image biological samples. A single protein-DNA binding reaction imaged by AFM can reveal protein binding specificity and affinity, protein-induced DNA bending, and protein binding stoichiometry. Changes in DNA structure, complex conformation, and cooperativity, can also be analyzed. In this review we highlight some important examples in the literature and discuss the advantages and limitations of these measurements. We also discuss important advances in technology that will facilitate the progress of AFM in the future.

Contents

Abbreviations: AFM, atomic force microscopy; AGT, O^6-alkylguanine DNA alkyltransferase; CPD, cyclobutane pyrimidine dimer; DREEM, dual-resonance-frequency-enhanced electrostatic force microscopy; EM, electron microscopy; EMSA, electrophoretic mobility shift assay; FRET, Förster resonance energy transfer; HS-AFM, high-speed AFM; MW, molecular weight; NER, nucleotide excision repair; NMR, nuclear magnetic resonance; OGG1, 8-oxoguanine DNA glycosylase; SSB, single-stranded DNA binding protein; TDG, thymine DNA glycosylase; TIRF-AFM, combined total internal reflection fluorescence and AFM; UV-DDB, UV-damaged DNA-binding protein.

* Corresponding author at: 5117 Centre Ave, Hillman Cancer Center, Research Pavilion, Suite 2.6, Pittsburgh, PA 15213, USA.
E-mail address: vanhoutenb@upmc.edu (B. Van Houten).

https://doi.org/10.1016/j.semcdb.2017.06.028
1084-9521/

1. Introduction and scope

Developed in the mid-1980′s, the atomic force microscope has become an increasingly powerful instrument for studying physical properties of materials on an atomic scale [1]. In this chapter, we discuss the topology-based applications of this tool for studying interactions between proteins and DNA. When studying protein-DNA interactions, it is important to keep in mind the forces that govern them. The four major interactions are: (1) hydrogen bonding between side chain and main chain amino acids and the floor of the major or minor groove of the DNA helix, (2) ionic interactions between the negatively charged phosphate backbone of DNA and positive amino acid side chains, (3) hydrophobic interactions and particularly pi stacking of DNA bases and aromatic side chains, and (4) Van der Waals forces over large surface areas [2,3].

The first reports of atomic force microscopy (AFM) imaging of protein-DNA complexes were in 1992 of the *E. coli* RNA polymerase complexed with DNA [4] and of DNA polymerase on M13 phage DNA [5]. Since then, AFM has proven to have unique advantages in the study of protein-nucleic acid interactions. AFM imaging is relatively simple and the process allows for samples to remain under more physiological conditions. Relatively long DNA substrates may be used and there is no requirement for labeling, staining, or fixation of either the DNA or the protein. Most importantly, it provides direct imaging at the single molecule level, and thus rare events can be observed that would otherwise by obscured in bulk biochemistry techniques.

AFM is used to obtain topographical data of a sample on an atomically smooth substrate. The three main operating modes are contact, non-contact, and tapping or oscillating. Tapping mode is the most common for protein-DNA and other biological studies as it minimizes sample disturbance on the substrate. A probe tip at the end of an oscillating cantilever scans the sample and allows for three-dimensional imaging [6]. All studies discussed in this review use AFM tapping mode in air, unless otherwise specified.

Three major measurements can be obtained from a single AFM protein-DNA experiment, each providing unique insight into how a protein recognizes its target DNA. In Section 2, we discuss how DNA binding position relates to protein specificity and affinity. In Section 3, we review how AFM is used to image protein-induced DNA bending. Section 4 covers how AFM volumes can be used to determine stoichiometry of proteins binding to DNA. In addition to these three measurements, AFM can be used to study complex conformation, changes in DNA structure, and cooperative binding (Section 5). Finally, we will discuss recent technological advances and modifications to standard AFM that have contributed significantly to the study of protein-DNA interactions (Section 6).

2. Binding position

2.1. Specificity

The first, and perhaps easiest piece of information that can be obtained from an AFM protein-DNA experiment is binding position. This is obtained by measuring the contour length of the DNA molecule from one end to the center of a bound protein (Fig. 1A). Binding position can be measured in pixels and then reported either as a percent of the total contour length of the DNA or converted to nm or bp based on the image resolution. Each experiment will generate position values for all individual binding events, which are typically reported in a histogram and fit to a Gaussian distribution.

Binding position can be directly related to a protein's specificity for a DNA target. Non-specific DNA-binding proteins, then, are expected to be found distributed randomly along a substrate. If a specific target site is added, however, one expects to find the protein bound at that position with greater frequency. The standard deviation of the Gaussian represents the broadness of the distribution, and is indicative of protein specificity for a particular site. Specificity (S) can also be quantified from such data:

$$S = N \times \frac{A_{sp}}{A_{nsp}} + 1 \quad (1)$$

where N is the total number of binding sites along the DNA substrate and A_{sp}/A_{nsp} is the ratio of specific to non-specific binding, as determined by the area under the Gaussian [7]. The accuracy of these measurements depends on the uniformity of lengths of the DNA population. Modifications, such as biotinylation [8] or looping [9], to a specific end of a DNA substrate can help determine its orientation.

AFM can be used to demonstrate specificity for defined DNA sequences (e.g. transcription factors) or site-specific modifications (e.g. DNA repair proteins). In one early study, binding of the transcription factor AP2 was quantitatively mapped to the Na^+/K^+-ATPase $\alpha 1$ promoter [10]. Another study confirmed specific binding of the Pho4 protein within the *PHO5* promoter and identified previously unknown binding sites for the Mig1 protein in the *HXK2* promoter [11]. AFM has also demonstrated specific binding of the base excision repair protein 8-oxoguanine DNA glycosylase (OGG1) to an 8-oxoguanine adduct [12]. Non-specific binding for this interaction was also observed, allowing for important conclusions to be made about the lesion search mechanism. Recent studies on the binding position of the helicase XPD, involved in damage verification during nucleotide excision repair (NER), have given new insight into the mechanism of lesion recognition. Tessmer and colleagues identified archaeal XPD [13], and then eukaryotic XPD with p44 [14], stalling at a lesion by measuring its position via AFM. By engineering specific DNA substrates with spatially separated lesions and unpaired bubbles at known positions, the authors were able to identify unique recognition strategies of XPD for a fluorescein compared to a cyclobutane pyrimidine dimer (CPD). Other important examples include (but are not limited to) bacteriophage Ø29 proteins p4 and p6 [15], bacteriophage λ repressor CI [16], SA1 and TRF1 (involved in sister telomere cohesion) [17], and Ver (part of the *Drosophila* telomere-associated terminin complex) [18]. A report by Rippe and colleagues demonstrated the importance of deposition protocol on specificity measurements; samples that were more aggressively washed showed an increase in the non-specific binding of RNA polymerase with σ^{54} to promoter DNA [19].

2.2. Equilibrium binding

The majority of methods used to study the thermodynamic properties of protein-DNA interactions, such as the electrophoretic mobility shift assay (EMSA), DNase I footprinting, filter binding, fluorescence anisotropy, and surface plasmon resonance, are bulk assays. While useful in many applications, drawbacks include the inability to distinguish specific binding from nonspecific binding to

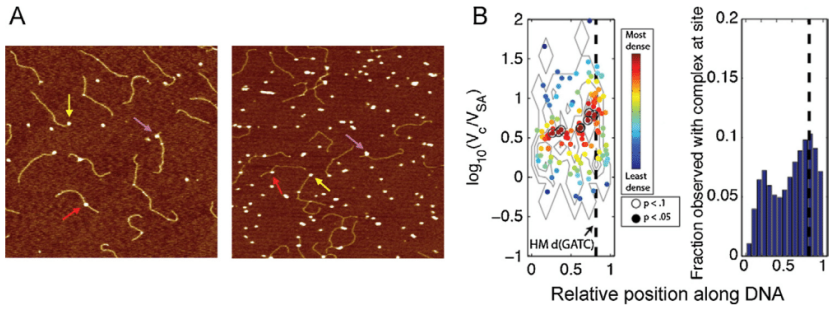

Fig. 1. Binding position of mismatch repair proteins on DNA. A. AFM images (1 μm × 1 μm, 2 nm height scale) of wildtype (left) and F39A mutant (right) MutS binding to 738-bp duplex DNA containing a T-bulge. Arrows indicate protein-DNA complexes at different DNA sites: DNA ends (yellow), specific complexes at a T-bulge (red), and nonspecific complexes (pink). Adapted with permission from [28]. B. Distribution of different complexes of MutS, MutL, and MutH on DNA containing a single hemi-methylated d(GATC) site (dashed line). Left, scatter plot of observed complex size vs. relative position on un-looped DNA molecules. Volumes are expressed as the complex volume (V_C) scaled to a streptavidin label (V_{SA}). Points are colored according to density scalebar; open circles are points that were more enriched than >90% of species of the indicated volume by permutation analysis. Right, distribution of observed occupancy of all protein complexes along DNA. Adapted with permission from [29].

DNA ends, a common property of DNA-binding proteins [7,20,21]. In 2005, Erie and colleagues showed that binding positions collected by AFM could also be used to directly determine protein-DNA equilibrium constants at the single molecule level [7]. Basing their calculations on the lattice binding model for protein-DNA interactions [22], they derived equations to calculate binding constants from AFM data. For the common experimental case, where DNA substrates are relatively long and contain one site-specific binding target, the equilibrium association constant for that site, K_{SP}, can be approximated as

$$K_{SP} \approx \frac{O_{SP}}{(1 - O_{SP}) \times ([P] - [D] \times O_{Fragment})} \quad (2)$$

where O_{SP} is the fractional occupancy at the target site, [P] and [D] are the total concentrations of protein and DNA substrate, and $O_{Fragment}$ is the average number of proteins bound per DNA molecule [7]. The equilibrium dissociation constant, K_D, would be the reciprocal of this equation.

This calculation is based on the assumption that equilibrium populations in solution are conserved during the deposition process. Differences in the deposition affinity of bound and free DNA may obscure the binding affinity. However, the authors demonstrate that in cases where protein occupancy on the DNA is low and protein binding does not induce a three-dimensional structural change in the DNA, AFM is a reliable method. As such, this approach has been applied to DNA-binding studies of MutS (discussed below, Section 2.3.), RNA polymerase [23], Cas9 [24], PARP1 and PARP2 [25], and GabR [26], among others.

2.3. Case study: MutS binding position

Specificity measurements using AFM have been applied to several studies of bacterial mismatch repair. Early steps of this pathway involve mismatch recognition and binding by MutS and subsequent recruitment of MutL and MutH. In an initial study by Erie and colleagues on MutS-induced DNA bending (see below, Section 3), binding position was measured as a percentage of the total DNA contour length and used to separate populations of specific and non-specific complexes [27]. MutS exhibited either a Gaussian distribution centered at a mismatch or a random distribution on homoduplex DNA. At the mismatch, MutS exhibited two conformations: bent and unbent.

Continuing studies on MutS, this group sought to quantify binding constants and specificities via AFM [7]. They looked at MutS binding to homoduplex DNA, DNA containing a GT mismatch, and DNA containing an unpaired T. Based on their AFM data, they calculated binding constants for specific (mismatch, unpaired T, or DNA ends) and nonspecific sites. In order to compare to bulk estimates, they also calculated macroscopic DNA binding constants; these were either very similar to those based on bulk studies, or exposed potential artifacts of binding affinity by EMSA [7]. They then used this method to demonstrate changes in binding affinity and specificity for two mutants of MutS (Fig. 1A) [28].

Marszalek and colleagues determined the stoichiometry (see Section 4) of MutS, MutL, and MutH complexes and mapped their positions along a DNA substrate (Fig. 1B) [29]. By analyzing complex size and composition together with binding position, they were able to visualize heretofore unseen details of the early stages of mismatch repair.

3. Protein-induced DNA bend angles and flexibility

3.1. Energetics of protein-induced DNA bending

DNA is essentially a dynamic and elastic rod, and can be studied using the worm-like chain model. Free dsDNA has a persistence length (L_P, a measure of polymer stiffness) of ~50 nm and a free energy cost of ~1.5 kJ per degree of bend [30–33]. This DNA stiffness depends partially on sequence, chemical modifications to the DNA, binding of a protein, or some combination thereof. Some sequences can cause either a static or dynamic bend. Certain DNA lesions can distort the double helix and cause site-specific DNA opening and bending, reviewed in [34]. Here, we focus on protein-related DNA bending.

There are two general mechanisms for protein-induced bending of DNA. The first mechanism of bending is entropy driven: positive charges on the protein neutralize the negatively charged DNA backbone and release counterions into solution [35,36]. This often applies to proteins that bind AT-rich regions of the minor groove of DNA, which, in solution, is hydrated with a dense and highly ordered array of water molecules conferring significant rigidity to the DNA molecule, called a spine of hydration [37,38]. Protein binding and release of these molecules results in a loss of overall stiffness and an increase in entropy [32]. In the second mechanism, protein side chains insert between DNA bases, resulting in sharp kinks. Many protein-induced bends are governed by a combination of these two forces. For further discussion on the energetics of DNA bending, see reviews [32,39,40].

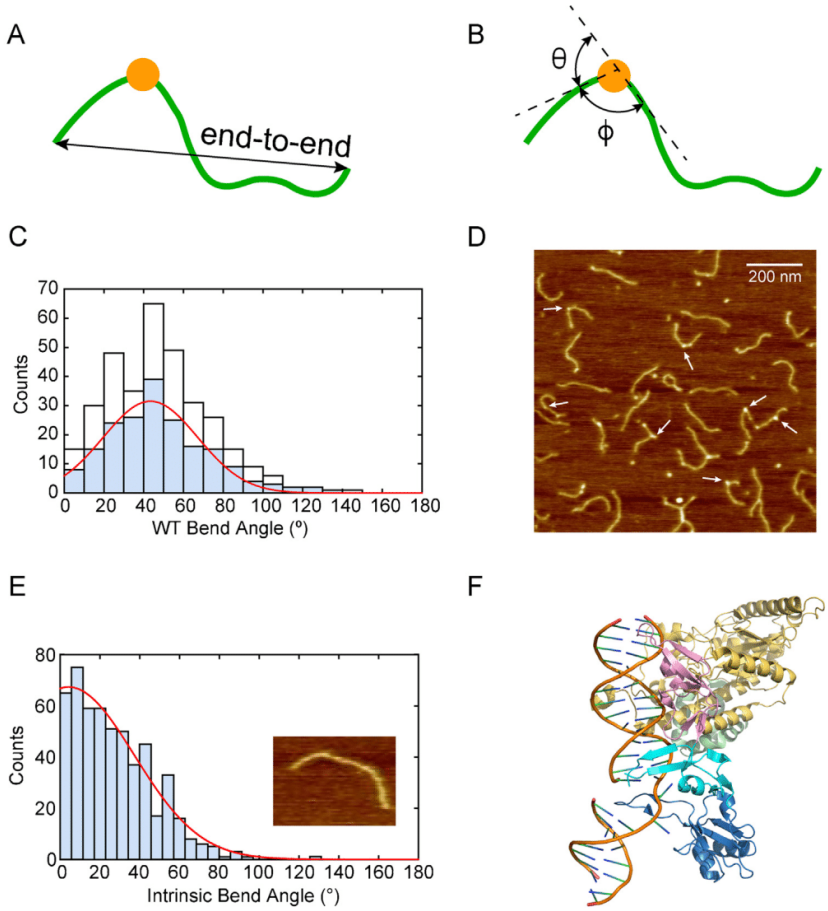

Fig. 2. DNA bending measurements and Rad4-Rad23-induced DNA bend angle. A. and B. Schematics showing the end-to-end measurement and the tangent method, respectively, for quantifying DNA bending. C. Distributions of Rad4-Rad23-induced DNA bend angles at all internal (white, n = 335) and specific (blue, n = 189) binding sites in the 538-bp fluorescein-dT containing DNA. Gaussian fitting (red curve) is shown for specific binding events only. D. AFM image of Rad4-Rad23 bound to the 538-bp fluorescein-dT containing DNA. White arrows highlight representative binding events. E. Distribution and Gaussian fitting of intrinsic DNA bend angles of the 538-bp DNA duplex containing a fluorescein-dT at 30% from one end (n = 245). Inset: representative AFM image of a 538-bp DNA fragment. C-E. Adapted with permission from [54]. F. Co-crystal structure of Rad4-Rad23 bound to DNA containing a CPD (PDB: 2QSG).

3.2. Methods to measure protein-induced bend angles

There are numerous methods that can be used to determine bend angles in DNA substrates, each with unique advantages and drawbacks. Structural methods like x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and electron microscopy (EM), have the potential to generate high resolution data for measuring protein-induced DNA bend angles. Biochemical methods include cyclization by ligase and gel mobility experiments, such as circular permutation. Biophysical methods include Förster resonance energy transfer (FRET) and optical or magnetic tweezers. Compared to the above methods, AFM offers some important benefits. AFM provides the most direct view of DNA bending with minimal sample manipulation. It also produces a low frequency of DNA bending artifacts if careful protocols are followed to avoid kinetic trapping during adsorption [31,33]. Finally, AFM bestows all the advantages of a single molecule technique, and thus distributions of bend angles can be analyzed for distinct populations or changes in stiffness.

3.2.1. Mean squared end-to-end distance

Assuming deposition maintains equilibrium conditions, the mean squared end-to-end distance, $\langle R^2 \rangle$, of the DNA polymer is:

$$\langle R^2 \rangle = 4L_P L_C \left(1 - \frac{2L_P}{L_C}\left(1 - e^{-L_C/2L_P}\right)\right) \quad (3)$$

Where L_P is the persistence length and L_C is the contour length of the DNA [31]. Contour length and end-to-end distances can easily be measured in AFM images (Fig. 2A) and used to determine the persistence length. In fact, these calculations can be used to test the assumption of two-dimensional equilibrium and verify that deposition has not altered the DNA conformation [19,41–43].

Protein-induced changes in DNA flexibility have been reported using this method for proteins including Rad50 [44], Abf2p [45], and HU [46]. Furthermore, end-to-end measurements can be used to estimate local DNA bend angles [47–50] with varying degrees of reliability compared to the tangent method (Section 3.2.2.).

3.2.2. Tangent method

The most direct way to measure DNA bending by AFM is called the tangent method. Vertices of interest (sites where protein is

Table 1
Comparison of protein-induced DNA bend angles (θ) determined by AFM (tangent method) and x-ray crystallography. AFM values reported as the mean and standard deviation (SD) of Gaussian fits to histogram data.

	ATOMIC FORCE MICROSCOPY				X-RAY CRYSTALLOGRAPHY			
Protein	DNA	Bend °	SD °	Ref.	DNA	Bend °	PDB	Ref.
BASE EXCISION REPAIR								
OGG1, *H. sapiens*	Nonspecific	0, 33, 70 [1]	9, 9, 9	[55]	Nonspecific	80	1YQK	[116] [2]
	Nonspecific	0, 70	21, 9.3	[12]	Tetrahydrofuran	70	1FN7	[117] [3]
	8-oxoguanine	71	9.2	[12] [4]	8-oxoguanine	70	1EBM	[118] [3,4]
TDG, *H. sapiens*	Nonspecific	33, 68	11, 11	[55]	Tetrahydrofuran	43	2RBA	[119] [5]
	Nonspecific	31, 65	12, 12	[55] [6]	5-carboxylcytosine	45	3UO7	[120] [5,6]
	G:U	29, 68	10, 10	[55] [6]				
	G:T	26, 65	15, 15	[55] [6]				
AlkA, *E. coli*	Nonspecific	0, 72	Not publ., 9	[12]	1-azaribose	66	1DIZ	[121]
DNA Pol β, *H. sapiens*	Nicked	39	34	Beckwitt et al, unpubl.	Nicked; 1 bp gap	90	1BPZ, 1BPX	[122]
MISMATCH REPAIR								
MutS, *E. coli*	Nonspecific	47–58 [7]	49–52	[27]	A:A; C:A; G:G	60	1OH5-7	[123] [5]
	Nonspecific	39–62	40–51	[27] [8]	Unpaired T	60	1OH8	[123] [5]
	Unpaired T	0, 32–42	10–11, 30–38	[27] [8]	Unpaired T	60	1EWQ	[124] [5,8]
	G:T	0, 74	12, 45	[27]	G:T	60	1E3M	[125] [5]
NUCLEOTIDE EXCISION REPAIR								
Rad4-Rad23, *S. cerevisiae*	Nonspecific	48.4	34.2	[53]	Nonspecific	42	4YIR	[53] [9]
	Nonspecific	49	34	[51] [10]	TT:TT	42	2QSH	[52] [5,11]
	Cholesterol	39	24	[51] [10]	CPD mismatch	42	2QSG	[52] [5,11]
	Fluorescein-dT	43	24	[54]				
DIRECT REVERSAL								
CPD photolyase, *A. nidulans*	Nonspecific	0	18	[126]				
	UV irrad.	36	30	[126]	CPD	50	1TEZ	[127]
AGT, *H. sapiens*	Nonspecific	0, 27, 58	13, 9, 29	[56]	N^4-alkylcytosine	15	1YFH	[128] [5]
CHROMATIN								
HU, *E. coli*	Nonspecific	Broad distrib.	N/A	[46]	Unpaired and mismatched T's	124	4QJU	[129] [12]
					Unpaired and mismatched T's	105–139	1P51, 1P71, 1P78	[130] [13]
IHF, *E. coli*	Consensus seq.	64	20	[47]	Consensus seq.	>160	1IHF	[131]
	Consensus seq.	123	Not publ.	[132]				
	Consensus seq.	50.36	26.67	[133] [14]				
Abf2p, *S. cerevisiae*	Nonspecific	78	Not publ.	[45]	AT-rich	90	5JGH, 5JH0	[134]
HMGB1 Box A+B, *R. norvegicus*	Nonspecific	67	21	[48]	AT-rich	85	4QR9	[135] [5,15]
TRANSCRIPTION								
Oct1, *H. sapiens*	Ad5 origin	42	12	[136] [5]	Octamer seq.	30	1OCT	[137] [5]
TtgV, *P. putida*	Operator	57	3	[138] [16]	Operator	60	2XRO	[139] [5,17]
Cro, Bacteriophage λ	Nonspecific	62	23	[140]	Operator	40	4CRO	[141]
	Operator	69	11	[140]	Operator	40	6CRO	[142]
DNA-MODIFYING ENZYMES								
M.HhaI, *H. haemolyticus*	Cognate seq.	2	28	[143]	Methylated 5-fluorocytosine	0	1MHT	[144]
Topoisomerase II, S. cerevisiae	4 bp overhangs	87	39	[49]	G-segment	150	2RGR	[145] [5,18]
Topoisomerase IIα, *H. sapiens*	4 bp overhangs	66	28	[49]	Doubly nicked	130	4FM9	[146] [5]

[1] Multiple values separated by commas indicate distinct populations observed in a single experiment.
[2] Mutant: N149C.
[3] Mutant: R9G, R10S, M11E.
[4] Mutant: K294Q.
[5] Truncated protein.
[6] Mutant: N140A.
[7] Range of values separated by a dash (–) indicates that multiple experiments were combined into one line.
[8] Organism: *T. aquaticus*.
[9] Mutant: K115T, V131C, C132S, V223E, Q427R.
[10] Organism: *H. sapiens* (homolog XPC-RAD23B).
[11] Mutant (Rad4): V223E, I225L.
[12] Organism: *S. aureus*.
[13] Organism: *Anabaena*.
[14] Organism: *B. abortus*.
[15] Second Box A domain instead of Box B.
[16] Imaged in liquid.
[17] Mutant: C109S, C205S.
[18] Mutant: P547L.

bound to DNA, for the purposes of this review) are identified in two-dimensional AFM images. Tangent lines following the path of the DNA immediately before and after the protein are drawn and the angle between them is called the opening angle, ϕ. The bend angle, θ, represents the degree that the DNA is bent from a straight line: θ = 180–ϕ (Fig. 2B).

Table 1 summarizes some important examples of DNA bend angles measured by the tangent method compared to those

reported from crystal structures. In many cases, such as for the NER complex Rad4-Rad23 (discussed further in Section 3.3.) [51–54], the two methods resulted in very similar bend angles. In other cases, as for OGG1 [12,55], thymine DNA glycosylase (TDG) [55], MutS [27], and O^6-alkylguanine DNA alkyltransferase (AGT) [56], AFM was able to detect multiple binding states that were apparently lost during crystal packing due to freezing out of one conformation. Furthermore, the width of the Gaussian fit to distributions of bend angles (SD in Table 1) is related to flexibility at that vertex. AFM is thus uniquely able to capture a range of steps in a dynamic bending process. It is interesting to note that differences in angle values may be due to variations in protein or DNA sequence or length, salt concentrations and buffer conditions, and overall flexibility of the DNA. Finally, one must consider that multiple protein binding and bending events on a single DNA molecule may not be co-planar and that the three-dimensional configuration of the bend angles when collapsed onto the two-dimensional mica surface can make analysis more difficult [33]. It is also useful to compare bend angles measured by AFM with single molecule FRET, as has been done for type IIA topoisomerases [49] and MutS [57].

3.3. Case study: Rad4-Rad23-induced DNA bend angle

In a recently published study on DNA damage recognition by yeast NER complex Rad4-Rad23, we showed that specific binding by the protein complex at fluorescein-dT sites in duplex DNA induced significant DNA bending ($43 \pm 24^\circ$, Fig. 2C,D), while the fluorescein-dT modification alone did not cause substantial bending in DNA ($4 \pm 32^\circ$, Fig. 2E) [54]. This bend angle induced by protein binding is in close agreement with that reported in a previously published crystal structure (42°, Fig. 2F) [52]. In addition, DNA was bent to a similar extent at nonspecific binding sites, consistent with an earlier study from the same group [53]. Significant DNA bending has also been reported for XPC-RAD23B (human homolog of Rad4-Rad23) binding to cholesterol-containing DNA [51].

4. Volume of protein-DNA complexes

4.1. Stoichiometry

The above two measurements (binding position and bend angle) rely on two-dimensional data. Peaks representing protein-DNA complexes in three-dimensional topological AFM images can be analyzed for volumes. Several methods for volume determination have been reviewed [58]. After identifying nucleoprotein complexes, one common method to calculate volume is to first determine the average height of the complex and subtract the average height for the whole image (i.e. background). This is multiplied by the area of the footprint of the complex to give approximate volumes (Fig. 3A).

In general, due to tip effects [59], deposition protocols, and sample deformation, sample heights and widths tend to be underestimated and overestimated, respectively, by AFM. Thus, AFM volumes may not reflect actual sample volumes, but are still directly proportional [58,60–62]. Schneider and colleagues demonstrated the linear relationship between AFM volumes and molecular weight (MW) using several globular proteins of known sizes [60]. Thus, they demonstrated the ability to distinguish between oligomeric states of proteins. A standard curve (such as the one in Fig. 3B) enables determination of MW and therefore stoichiometry of unknown proteins. Such standards, however, are not universal, and only apply to individual AFM set-ups using the same tip and deposition/imaging methods.

The same principles should apply for proteins bound to DNA. For complexes with relatively large proteins, the contribution of the DNA volume does not significantly increase the total volume of the complex. This method has identified stoichiometries on DNA for MutS [27,29,63], NER proteins UvrB/C and XPD/p44 [14], HIV restriction factor A3G [64,65], GabR [26], Ver [18], and many others. Some groups subtract out the volume of the DNA [66,67], which becomes particularly important when protein volumes are relatively small.

4.2. Case study: UV-DDB stoichiometry on DNA

In one thorough study of UV-damaged DNA-binding protein (UV-DDB), AFM was used to confirm and directly visualize oligomerization on damaged DNA. UV-DDB is an important protein involved in lesion recognition during global genome NER of UV-induced photoproducts in the context of chromatin [68]. It is known that UV-DDB is made up of two proteins, DDB1 (127 kDa) and DDB2 (48 kDa) [69], but its stoichiometry on DNA was unclear. Yeh and colleagues published a crystal structure showing that full-length human UV-DDB binds to two DNA molecules, each containing a tetrahydrofuran lesion (a stable mimic of an abasic site) as a dimer of dimers ($(DDB1\text{-}DDB2)_2$, Fig. 3C) [70]. This was further supported by protein surface area calculations of negative-stain EM as well as dynamic light scattering experiments [70].

In the above study, AFM was used as an important complement to these methods, providing more physiological conditions (longer DNA substrates and more physiological ionic strength) and thus able to support the biological relevance of the dimeric structure. Binding reactions of UV-DDB with UV-irradiated 517 bp dsDNA fragments were deposited on mica and imaged using tapping AFM [70]. UV-irradiation generates CPDs and (6-4) photoproducts, which are canonical substrates for NER. AFM experiments, then, were able to detect UV-DDB binding to either one or two DNA molecules simultaneously (Fig. 3D). Volume analysis of these two states revealed mean volumes of 507 or 530 nm^3 (Fig. 3E), corresponding to 350 or 365 kDa, respectively (Fig. 3C). Again, this is consistent with the dimer of dimers model. Similar AFM studies demonstrated that a xeroderma pigmentosum-causing mutant of DDB2, K244E, lost all binding specificity but was still able to form a dimer of dimers (Fig. 3F,G) [21].

5. Other AFM measurements for protein-DNA interactions

5.1. Changes in DNA structure

In addition to bending, other protein-induced structural changes to DNA can be observed with AFM. Changes in contour length of the DNA can signify compaction and deviation from the B-form helix or wrapping in a nucleoprotein complex. DNA looping can also be observed.

RNA polymerase has been studied extensively by AFM to demonstrate changes in DNA structure [19,23,42,71–73]. Bustamante and colleagues provided evidence that binding of *E. coli* RNA polymerase-σ^{54} and NtrC (constitutive mutant D54E, S160F) was able to induce DNA looping; they obtained images of multiple intermediates during activation of transcription [71]. Several groups have reported a decrease in DNA contour length upon RNA polymerase binding, indicating that the DNA is making more contacts with the protein than would be expected for a linear path (i.e. wrapped in the complex) [23,42,72]. Rippe and colleagues also observed a reduction in DNA contour length upon binding of RNA polymerase-σ^{54}, signifying wrapping in the complex, but warned that this may be partly an artifact of the deposition and drying process [19]. Koroleva and colleagues were able to distinguish between

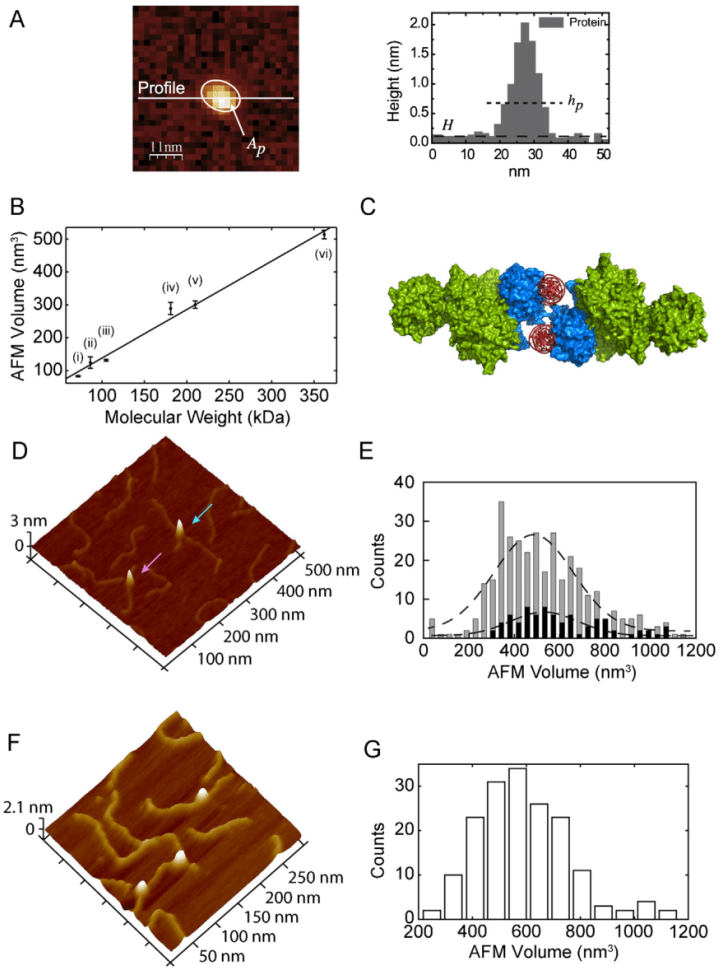

Fig. 3. Protein volume and stoichiometry of UV-DDB. A. Schematic showing one method of measurement of AFM volume of a protein particle. Image on right is a cross-section showing heights of image on left. The area, A_p, of a slice is taken at a certain height, h_p, above the background, H. The average height of the particle, h_{avg}, is determined (not shown). $Volume = A_p\,(h_{avg} - H)$. Adapted with permission from [58]. B. Calibration curve relating the molecular weight of a complex to its measured AFM volume, mean ± SD of three separate determinations. The curve was generated using the following proteins in solution: (i) Pot1 monomer (65 kDa), (ii) PcrA monomer (86.4 kDa), (iii) UvrA monomer (105 kDa), (iv) Taq MutS dimer (181 kDa), (v) UvrA dimer (210 kDa), and (vi) Taq MutS tetramer (362 kDa). Linear fit to the data yields $V(nm^3)$=1.471MW(kDa)-7.294 with R^2=0.9886. C. Crystal structure of UV-DDB in complex with damaged DNA, showing dimer of dimers in surface representation (PDB: 4E5Z). D. AFM image of wildtype UV-DDB binding to UV-irradiated 517-bp duplex DNA. Pink and blue arrows indicate dimeric UV-DDB binding to one and two molecules of duplex DNA, respectively. E. Distributions and Gaussian fittings (dashed lines) of AFM volume of UV-DDB on one (gray bars, n = 339) and two strands (black bars, n = 79) of duplex DNA. F. AFM image of UV-DDB (K244E) bound to 517-bp duplex DNA as dimer of dimers. G. Distribution of AFM volume of UV-DDB (K244E) on DNA (n = 171). B, F, G. Adapted with permission from [21]. D, E. Adapted with permission from [70].

cis and *trans* configurations of two RNA polymerase proteins on a single DNA substrate [73].

Protein-induced reduction of DNA contour length has identified DNA wrapping in several other protein-DNA complexes, including the centromeric DNA-binding factor 3 bound to the kinetochore [74], the LrpC protein bound to *lrpC* promoter DNA [75], single-stranded DNA binding protein (SSB) bound to ssDNA [76], and multifunctional aminopeptidase A bound to the *carAB* operon [43]. In other studies, small reductions in DNA contour length have been interpreted as overall compaction, such as for nuclear factor I at the Ad5 origin [77]. The mitochondrial transcription factor TFAM is able to loop, organize, and compact plasmid DNA; stages of compaction were observed by AFM at increasing protein concentrations by non-contact AFM [78]. In a final example, the bacteriophage λ repressor CI was seen oligomerizing on and inducing looping of operator DNA sequences [16].

5.2. Complex conformation

Although not as high resolution as methods such as EM or x-ray crystallography, AFM can be used to visualize overall structural shapes in relatively physiological conditions. Differences in protein sizes (and thus AFM volumes) can be used to identify specific proteins and their arrangement within a complex. AFM has been used to identify the unique complex architecture of human Rad50 with Mre11 on DNA [44]. It has also captured complexes of the bacteriophage Ø29 proteins p4 and p6 in complex with Ø29 DNA and host *B. subtilis* RNA polymerase-σ^A; protein binding induces a distinct nucleoprotein-hairpin [15,79]. Furthermore, conformational

changes of MutL alpha [80] and condensin Smc2-Smc4 dimers [81] have been directly visualized.

5.3. Cooperativity

Particularly in the case of nucleoprotein filaments, AFM can be useful in demonstrating cooperative binding. By definition, proteins that bind DNA with high levels of cooperativity are more likely to bind to DNA that already has protein on it. This can be directly visualized by comparing the frequency of "saturated" DNA substrates to singly or partially bound substrates.

Cooperative binding by Mlh1-Pms1, involved in yeast mismatch repair, was reported using a combination of nitrocellulose filter binding assays and AFM [82]. At the single molecule level, the authors observed long continuous tracts of bound protein at a much higher frequency than singly bound heterodimers. In another study, the *E. coli* nucleoid-associated protein HU was shown to both bend (Table 1) and condense DNA at low concentrations, but formed a rigid nucleoprotein filament at high concentrations [46]. Filament formation appeared highly cooperative because increasing protein concentration caused a shift from free or partially bound DNA to 100% rigid nucleoprotein filaments. Other examples include RdgC [83], RecA and SSB [76,84], and AGT [56].

Similarly to binding affinity (Section 2.2.), cooperativity can be quantified based on the lattice model [22,85]. Dekker and colleagues applied this theory to experimental AFM data with maximum-likelihood data analysis and published the following analytical solution for the degree of cooperativity, ω:

$$\omega = \frac{S\theta nL + \theta S^2 - \theta nL^2 - \theta SL\text{-}nSL\text{+}nL^2}{\theta S^2} \quad (4)$$

where S = total number of data points, θ = DNA saturation (bound contour length/total contour length for all DNA molecules), n = size of the binding site in base pairs or nucleotide residues, $L = \sum_{i=1}^{S} c_i$, and c = filament length (c_i = the number of contiguously bound proteins in the ith filament) [86]. Positive cooperativity is signified by ω values greater than 1. They found that topoisomerase IB from vaccinia virus bound DNA with high levels of cooperativity; $\omega = 5.6\text{–}9.1 \times 10^3$ [86].

6. New AFM tools to study protein-DNA interactions

6.1. Time-lapse and high-speed AFM

Development of advanced AFM modes and combination with other techniques have created even more powerful tools in the field of molecular and biological imaging [87]. Soon after the development of the atomic force microscope, Bustamante, Thomson, and colleagues showed that RNA polymerase could transcribe DNA while adsorbed on mica. They applied time-lapse imaging techniques to capture sequential images following the assembly of RNA polymerase-DNA complexes and polymerase sliding along the DNA in liquid [88–91]. However, the significant difference in timescales for standard AFM tapping mode (minutes per image) and biological processes (seconds or less) limited their ability to detect these dynamic events; faster scanning was required. The first high-speed AFM (HS-AFM) suitable for biological imaging was established by Ando and colleagues in 2001 [92]. See reference [93] for a nice review on its development. By using smaller cantilevers with an optical beam deflection detector, high feedback bandwidth, and a high-speed scanner, they were able to obtain multiple frames per second. Lyubchenko and colleagues have used HS-AFM to provide unprecedented insight into DNA wrapping and unwrapping around nucleosomes [94,95], and dynamic A3G complex formation [96,97].

6.2. Fluorescence-coupled AFM

Particularly when using AFM to analyze complex conformations of protein and DNA, nondescript globular structures can make it difficult to identify components. Combining AFM with fluorescence microscopy addresses this issue and enables the study of multiple proteins on DNA [98–100].

Investigations of homologous recombination involved in double-strand break repair by Wyman and colleagues have established a combination of atomic force and total internal reflection fluorescence microscopy (TIRF-AFM) and demonstrated the usefulness of its application toward protein-DNA studies. Fluorescently labeled RAD54 was localized on RAD51-DNA filaments and distinct conformations were visualized that would otherwise be undistinguishable due to similar overall dimensions [101]. The same group recently published TIRF-AFM results following the BRCA2 hand-off to RAD51 [102].

6.3. Dual-resonance-frequency-enhanced electrostatic force microscopy

While fluorescence-coupled AFM methods seek to localize specific proteins within a complex, dual-resonance-frequency-enhanced electrostatic force microscopy (DREEM) enables resolution of DNA within a complex [103]. The union of topography and DREEM phase imaging, developed by Wang, Erie, and colleagues, has enabled high resolution visualization of the path of dsDNA around a nucleosome [103], ssDNA wrapping around nucleosome-like histone complexes [104] and the organization of telomeric DNA with TRF2 [105,106]. For more information on electrostatic force microscopy and multiparametric AFM, see references [107,108].

6.4. New modes of AFM operation

The advent of tapping AFM [6] was a great turning point for biological research. Since then, modifications to this mode have further improved the field in terms of resolution and sample disturbance. PeakForce, a proprietary mode developed by Bruker, is similar to tapping mode, but uses a slower oscillating cantilever and a feedback loop to finely control the tip interaction with the sample. Williams and colleagues used this mode to make many of the measurements discussed above; they determined the binding specificity, DNA bend angle and flexibility, and DNA compaction induced by HMO1 binding [109]. We have used PeakForce to study DNA bending by DNA polymerase β (Table 1, Beckwitt et al., unpublished). It has also been applied to studies of nucleosome structure [110], RecG and SSB [111], BRCA1 [112], and Cox [113].

A similar effect is achieved by small amplitude small set-point (SASS). By limiting the tip oscillation amplitude to within the surface water films, noise is reduced and resolution is enhanced [114]. Maxwell and colleagues demonstrated an improvement in image quality in a report on type II topoisomerase-induced DNA bending [115].

7. Outlook

AFM holds great promise for the study of protein-DNA interactions both statically on dried surfaces and dynamically in solution in real-time. The key features of this tool allow for label free analysis of populations of molecules that show amazing diversity in their structures and conformations, and thus fills an important scientific niche complementing x-ray crystallography, NMR, and

cryo-EM. Analysis of the dynamics of protein-DNA complex formation involving multiple proteins and unique hand-offs from one protein to the next using high speed AFM will provide a unique view of protein kinetics. Finally, future studies combining multiple platforms, as described above, will certainly enhance our understanding of how proteins interact with DNA to recognize their cognate binding sites.

Acknowledgements

This work is supported by funding from the National Institutes of Health 5R01ES019566 (B.V.H.) and T32GM088119 (E.C.B.). The authors thank Hong Wang for careful reading of the manuscript and thoughtful feedback. Due to space constraints, we regret that many important original papers could not be directly cited and discussed.

References

[1] G. Binnig, C.F. Quate, C. Gerber, Atomic force microscope, Phys. Rev. Lett. 56 (9) (1986) 930–933.
[2] R.T. Sauer, Protein-DNA Interactions, Methods in Enzymology, Academic Press, San Diego, California, 1991.
[3] L. Jen-Jacobson, L.E. Engler, L.A. Jacobson, Structural and thermodynamic strategies for site-specific DNA binding proteins, Structure 8 (10) (2000) 1015–1023.
[4] C. Bustamante, J. Vesenka, C.L. Tang, W. Rees, M. Guthold, R. Keller, Circular DNA molecules imaged in air by scanning force microscopy, Biochemistry 31 (1) (1992) 22–26.
[5] J. Yang, K. Takeyasu, Z. Shao, Atomic force microscopy of DNA molecules, FEBS Lett. 301 (2) (1992) 173–176.
[6] Q. Zhong, D. Inniss, K. Kjoller, V. Elings, Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy, Surf. Sci. 290 (1–2) (1993) L688–L692.
[7] Y. Yang, L.E. Sass, C. Du, P. Hsieh, D.A. Erie, Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: mutS-DNA interactions, Nucl. Acids Res. 33 (13) (2005) 4322–4334.
[8] Z. Voros, Y. Yan, D.T. Kovari, L. Finzi, D. Dunlap, Proteins mediating DNA loops effectively block transcription, Protein Sci. 26 (7) (2017) 1427–1438.
[9] O. Chammas, D.J. Billingsley, W.A. Bonass, N.H. Thomson, Single-stranded DNA loops as fiducial markers for exploring DNA-protein interactions in single molecule imaging, Methods 60 (2) (2013) 122–130.
[10] S. Nettikadan, F. Tokumasu, K. Takeyasu, Quantitative analysis of the transcription factor AP2 binding to DNA by atomic force microscopy, Biochem. Biophys. Res. Commun. 226 (3) (1996) 645–649.
[11] F. Moreno-Herrero, P. Herrero, J. Colchero, A.M. Baro, F. Moreno, Imaging and mapping protein-binding sites on DNA regulatory regions with atomic force microscopy, Biochem. Biophys. Res. Commun. 280 (1) (2001) 151–157.
[12] L. Chen, K.A. Haushalter, C.M. Lieber, G.L. Verdine, Direct visualization of a DNA glycosylase searching for damage, Chem. Biol. 9 (3) (2002) 345–350.
[13] C.N. Buechner, K. Heil, G. Michels, T. Carell, C. Kisker, I. Tessmer, Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility, J. Biol. Chem. 289 (6) (2014) 3613–3624.
[14] N. Wirth, J. Gross, H.M. Roth, C.N. Buechner, C. Kisker, I. Tessmer, Conservation and divergence in nucleotide excision repair lesion recognition, J. Biol. Chem. (2016).
[15] P.G. del Arroyo, M. Velez, O. Pietrement, M. Salas, J.L. Carrascosa, A. Camacho, A nucleoprotein-hairpin in transcription regulation, J. Struct. Biol. 168 (3) (2009) 444–451.
[16] H. Wang, L. Finzi, D.E. Lewis, D. Dunlap, AFM studies of lambda repressor oligomers securing DNA loops, Curr. Pharm. Biotechnol. 10 (5) (2009) 494–501.
[17] J. Lin, P. Countryman, H. Chen, H. Pan, Y. Fan, Y. Jiang, P. Kaur, W. Miao, G. Gurgel, C. You, J. Piehler, N.M. Kad, R. Riehn, P.L. Opresko, S. Smith, Y.J. Tao, H. Wang, Functional interplay between SA1 and TRF1 in telomeric DNA binding and DNA–DNA pairing, Nucleic Acids Res. 44 (13) (2016) 6363–6376.
[18] A. Cicconi, E. Micheli, F. Verni, A. Jackson, A.C. Gradilla, F. Cipressa, D. Raimondo, G. Bosso, J.G. Wakefield, L. Ciapponi, G. Cenci, M. Gatti, S. Cacchione, G.D. Raffa, The Drosophila telomere-capping protein Verrocchio binds single-stranded DNA and protects telomeres from DNA damage response, Nucl. Acids Res. (2016).
[19] A. Schulz, N. Mucke, J. Langowski, K. Rippe, Scanning force microscopy of Escherichia coli RNA polymerase.sigma54 holoenzyme complexes with DNA in buffer and in air, J. Mol. Biol. 283 (4) (1998) 821–836.
[20] H. Wang, I. Tessmer, D.L. Croteau, D.A. Erie, B. Van Houten, Functional characterization and atomic force microscopy of a DNA repair protein conjugated to a quantum dot, Nano Lett. 8 (6) (2008) 1631–1637.
[21] H. Ghodke, H. Wang, C.L. Hsieh, S. Woldemeskel, S.C. Watkins, V. Rapic-Otrin, B. Van Houten, Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates, Proc. Natl. Acad. Sci. U. S. A. 111 (18) (2014) E1862–71.
[22] J.D. McGhee, P.H. von Hippel, Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice, J. Mol. Biol. 86 (2) (1974) 469–489.
[23] N. Doniselli, P. Rodriguez-Aliaga, D. Amidani, J.A. Bardales, C. Bustamante, D.G. Guerra, C. Rivetti, New insights into the regulatory mechanisms of ppGpp and DksA on Escherichia coli RNA polymerase-promoter complex, Nucl. Acids Res. 43 (10) (2015) 5249–5262.
[24] E.A. Josephs, D.D. Kocak, C.J. Fitzgibbon, J. McMenemy, C.A. Gersbach, P.E. Marszalek, Structure and specificity of the RNA-guided endonuclease Cas9 during DNA interrogation, target binding and cleavage, Nucl. Acids Res. 43 (18) (2015) 8924–8941.
[25] M.V. Sukhanova, S. Abrakhi, V. Joshi, D. Pastre, M.M. Kutuzov, R.O. Anarbaev, P.A. Curmi, L. Hamon, O.I. Lavrik, Single molecule detection of PARP1 and PARP2 interaction with DNA strand breaks and their poly(ADP-ribosyl)ation using high-resolution AFM imaging, Nucl. Acids Res. 44 (6) (2016) e60.
[26] D. Amidani, A. Tramonti, A.V. Canosa, B. Campanini, S. Maggi, T. Milano, M.L. di Salvo, S. Pascarella, R. Contestabile, S. Bettati, C. Rivetti, Study of DNA binding and bending by Bacillus subtilis GabR, a PLP-dependent transcription factor, Biochim. Biophys. Acta 1861 (1 Pt A) (2017) 3474–3489.
[27] H. Wang, Y. Yang, M.J. Schofield, C. Du, Y. Fridman, S.D. Lee, E.D. Larson, J.T. Drummond, E. Alani, P. Hsieh, D.A. Erie, DNA bending and unbending by MutS govern mismatch recognition and specificity, Proc. Natl. Acad. Sci. U. S. A. 100 (25) (2003) 14822–14827.
[28] I. Tessmer, Y. Yang, J. Zhai, C. Du, P. Hsieh, M.M. Hingorani, D.A. Erie, Mechanism of MutS searching for DNA mismatches and signaling repair, J. Biol. Chem. 283 (52) (2008) 36646–36654.
[29] E.A. Josephs, T. Zheng, P.E. Marszalek, Atomic force microscopy captures the initiation of methyl-directed DNA mismatch repair, DNA Repair (Amst.) 35 (2015) 71–84.
[30] C. Bustamante, J.F. Marko, E.D. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA, Science 265 (5178) (1994) 1599–1600.
[31] C. Rivetti, M. Guthold, C. Bustamante, Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis, J. Mol. Biol. 264 (5) (1996) 919–932.
[32] P.L. Privalov, A.I. Dragan, C. Crane-Robinson, The cost of DNA bending, Trends Biochem. Sci. 34 (9) (2009) 464–470.
[33] C. Bustamante, C. Rivetti, Visualizing protein-nucleic acid interactions on a large scale with the scanning force microscope, Annu. Rev. Biophys. Biomol. Struct. 25 (1996) 395–429.
[34] B. Van Houten, Nucleotide excision repair in Escherichia coli, Microbiol. Rev. 54 (1) (1990) 18–51.
[35] G.S. Manning, The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides, Q. Rev. Biophys. 11 (2) (1978) 179–246.
[36] T.M. Lohman, D.P. Mascotti, Thermodynamics of ligand-nucleic acid interactions, Methods Enzymol. 212 (1992) 400–424.
[37] A.I. Dragan, D.J. Russell, P.L. Privalov, DNA hydration studied by pressure perturbation scanning microcalorimetry, Biopolymers 91 (1) (2009) 95–101.
[38] H.R. Drew, R.E. Dickerson, Structure of a B-DNA dodecamer. III. Geometry of hydration, J. Mol. Biol. 151 (3) (1981) 535–556.
[39] A. Ansari, S.V. Kuznetsov, Dynamics and mechanism of DNA-Bending proteins in binding site recognition, in: M.C. Williams, L.J. Maher (Eds.), Biophysics of DNA-Protein Interactions: From Single Molecules to Biological Systems, Springer New York, New York, NY, 2011, pp. 107–142.
[40] S. Harteis, S. Schneider, Making the bend: DNA tertiary structure and protein-DNA interactions, Int. J. Mol. Sci. 15 (7) (2014) 12335–12363.
[41] E. Margeat, C. Le Grimellec, C.A. Royer, Visualization of trp repressor and its complexes with DNA by atomic force microscopy, Biophys. J. 75 (6) (1998) 2712–2720.
[42] C. Rivetti, M. Guthold, C. Bustamante, Wrapping of DNA around the E.coli RNA polymerase open promoter complex, EMBO J. 18 (16) (1999) 4464–4475.
[43] P.N. Minh, N. Devroede, J. Massant, D. Maes, D. Charlier, Insights into the architecture and stoichiometry of Escherichia coli PepA*DNA complexes involved in transcriptional control and site-specific DNA recombination by atomic force microscopy, Nucl. Acids Res. 37 (5) (2009) 1463–1476.
[44] M. de Jager, J. van Noort, D.C. van Gent, C. Dekker, R. Kanaar, C. Wyman, Human Rad50/Mre11 is a flexible complex that can tether DNA ends, Mol. Cell 8 (5) (2001) 1129–1135.
[45] R.W. Friddle, J.E. Klare, S.S. Martin, M. Corzett, R. Balhorn, E.P. Baldwin, R.J. Baskin, A. Noy, Mechanism of DNA compaction by yeast mitochondrial protein Abf2p, Biophys. J. 86 (3) (2004) 1632–1639.
[46] J. van Noort, S. Verbrugge, N. Goosen, C. Dekker, R.T. Dame, Dual architectural roles of HU: formation of flexible hinges and rigid filaments, Proc. Natl. Acad. Sci. U. S. A. 101 (18) (2004) 6969–6974.
[47] R.T. Dame, J. van Mameren, M.S. Luijsterburg, M.E. Mysiak, A. Janicijevic, G. Pazdzior, P.C. van der Vliet, C. Wyman, G.J. Wuite, Analysis of scanning force microscopy images of protein-induced DNA bending using simulations, Nucl. Acids Res. 33 (7) (2005) e68.
[48] J. Zhang, M.J. McCauley, L.J. Maher 3rd, M.C. Williams, N.E. Israeloff, Mechanism of DNA flexibility enhancement by HMGB proteins, Nucl. Acids Res. 37 (4) (2009) 1107–1114.
[49] A.H. Hardin, S.K. Sarkar, Y. Seol, G.F. Liou, N. Osheroff, K.C. Neuman, Direct measurement of DNA bending by type IIA topoisomerases: implications for

non-equilibrium topology simplification, Nucl. Acids Res. 39 (13) (2011) 5729–5743.
[50] J. Zhang, M.J. McCauley, L.J. Maher 3rd, M.C. Williams, N.E. Israeloff, Basic N-terminus of yeast Nhp6A regulates the mechanism of its DNA flexibility enhancement, J. Mol. Biol. 416 (1) (2012) 10–20.
[51] A. Janicijevic, K. Sugasawa, Y. Shimizu, F. Hanaoka, N. Wijgers, M. Djurica, J.H. Hoeijmakers, C. Wyman, DNA bending by the human damage recognition complex XPC-HR23B, DNA Repair (Amst.) 2 (3) (2003) 325–336.
[52] J.H. Min, N.P. Pavletich, Recognition of DNA damage by the Rad4 nucleotide excision repair protein, Nature 449 (7162) (2007) 570–575.
[53] X. Chen, Y. Velmurugu, G. Zheng, B. Park, Y. Shim, Y. Kim, L. Liu, B. Van Houten, C. He, A. Ansari, J.H. Min, Kinetic gating mechanism of DNA damage recognition by Rad4/XPC, Nat. Commun. 6 (2015) 5849.
[54] M. Kong, L. Liu, X. Chen, K.I. Driscoll, P. Mao, S. Bohm, N.M. Kad, S.C. Watkins, K.A. Bernstein, J.J. Wyrick, J.H. Min, B. Van Houten, Single-molecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process, Mol. Cell 64 (2) (2016) 376–387.
[55] C.N. Buechner, A. Maiti, A.C. Drohat, I. Tessmer, Lesion search and recognition by thymine DNA glycosylase revealed by single molecule imaging, Nucl. Acids Res. 43 (5) (2015) 2716–2729.
[56] I. Tessmer, M. Melikishvili, M.G. Fried, Cooperative cluster formation, DNA bending and base-flipping by O6-alkylguanine-DNA alkyltransferase, Nucl. Acids Res. 40 (17) (2012) 8296–8308.
[57] V.C. DeRocco, L.E. Sass, R. Qiu, K.R. Weninger, D.A. Erie, Dynamics of MutS-mismatched DNA complexes are predictive of their repair phenotypes, Biochemistry 53 (12) (2014) 2043–2052.
[58] M.E. Fuentes-Perez, M.S. Dillingham, F. Moreno-Herrero, AFM volumetric methods for the characterization of proteins and nucleic acids, Methods 60 (2) (2013) 113–121.
[59] J. Vesenka, M. Guthold, C. Tang, D. Keller, E. Delaine, C. Bustamante, Substrate preparation for reliable imaging of DNA molecules with the scanning force microscope, Ultramicroscopy 42 (1992) 1243–1249.
[60] S.W. Schneider, J. Larmer, R.M. Henderson, H. Oberleithner, Molecular weights of individual proteins correlate with molecular volumes measured by atomic force microscopy, Pflugers Arch. 435 (3) (1998) 362–367.
[61] Y. Yang, H. Wang, D.A. Erie, Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy, Methods 29 (2) (2003) 175–187.
[62] G.C. Ratcliff, D.A. Erie, A novel single-molecule study to determine protein-protein association constants, J. Am. Chem. Soc. 123 (24) (2001) 5632–5635.
[63] R. Qiu, M. Sakato, E.J. Sacho, H. Wilkins, X. Zhang, P. Modrich, M.M. Hingorani, D.A. Erie, K.R. Weninger, MutL traps MutS at a DNA mismatch, Proc. Natl. Acad. Sci. U. S. A. 112 (35) (2015) 10914–10919.
[64] L. Chelico, E.J. Sacho, D.A. Erie, M.F. Goodman, A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV, J. Biol. Chem. 283 (20) (2008) 13780–13791.
[65] L.S. Shlyakhtenko, A.Y. Lushnikov, M. Li, L. Lackey, R.S. Harris, Y.L. Lyubchenko, Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G, J. Biol. Chem. 286 (5) (2011) 3387–3395.
[66] F. Pratto, Y. Suzuki, K. Takeyasu, J.C. Alonso, Single-molecule analysis of protein-DNA complexes formed during partition of newly replicated plasmid molecules in Streptococcus pyogenes, J. Biol. Chem. 284 (44) (2009) 30298–30306.
[67] C. Wyman, I. Rombel, A.K. North, C. Bustamante, S. Kustu, Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein, Science 275 (5306) (1997) 1658–1661.
[68] B. Van Houten, M. Kong, Eukaryotic nucleotide excision repair, in: R.A. Bradshaw, P.D. Stahl (Eds.), Encyclopedia of Cell Biology, Academic Press, 2016, pp. 435–441.
[69] R. Dualan, T. Brody, S. Keeney, A.F. Nichols, A. Admon, S. Linn, Chromosomal localization and cDNA cloning of the genes (DDB1 and DDB2) for the p127 and p48 subunits of a human damage-specific DNA binding protein, Genomics 29 (1) (1995) 62–69.
[70] J.I. Yeh, A.S. Levine, S. Du, U. Chinte, H. Ghodke, H. Wang, H. Shi, C.L. Hsieh, J.F. Conway, B. Van Houten, V. Rapic-Otrin, Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair, Proc. Natl. Acad. Sci. U. S. A. 109 (41) (2012) E2737–46.
[71] K. Rippe, M. Guthold, P.H. von Hippel, C. Bustamante, Transcriptional activation via DNA-looping: visualization of intermediates in the activation pathway of E. coli RNA polymerase x sigma 54 holoenzyme by scanning force microscopy, J. Mol. Biol. 270 (2) (1997) 125–138.
[72] C. Rivetti, S. Codeluppi, G. Dieci, C. Bustamante, Visualizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases, J. Mol. Biol. 326 (5) (2003) 1413–1426.
[73] O.N. Koroleva, E.V. Dubrovin, I.V. Yaminsky, V.L. Drutsa, Effect of DNA bending on transcriptional interference in the systems of closely spaced convergent promoters, Biochim. Biophys. Acta 1860 (10) (2016) 2086–2096.
[74] L.I. Pietrasanta, D. Thrower, W. Hsieh, S. Rao, O. Stemmann, J. Lechner, J. Carbon, H. Hansma, Probing the Saccharomyces cerevisiae centromeric DNA (CEN DNA)-binding factor 3 (CBF3) kinetochore complex by using atomic force microscopy, Proc. Natl. Acad. Sci. U. S. A. 96 (7) (1999) 3757–3762.
[75] C. Beloin, J. Jeusset, B. Revet, G. Mirambeau, F. Le Hegarat, E. Le Cam, Contribution of DNA conformation and topology in right-handed DNA wrapping by the Bacillus subtilis LrpC protein, J. Biol. Chem. 278 (7) (2003) 5333–5342.
[76] L. Hamon, D. Pastre, P. Dupaigne, C. Le Breton, E. Le Cam, O. Pietrement, High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein–DNA complexes, Nucleic Acids Res. 35 (8) (2007) e58.
[77] M.E. Mysiak, M.H. Bleijenberg, C. Wyman, P.E. Holthuizen, P.C. van der Vliet, Bending of adenovirus origin DNA by nuclear factor I as shown by scanning force microscopy is required for optimal DNA replication, J. Virol. 78 (4) (2004) 1928–1935.
[78] B.A. Kaufman, N. Durisic, J.M. Mativetsky, S. Costantino, M.A. Hancock, P. Grutter, E.A. Shoubridge, The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures, Mol. Biol. Cell 18 (9) (2007) 3225–3236.
[79] A. Camacho, M. Salas, Molecular interactions and protein-induced DNA hairpin in the transcriptional control of bacteriophage o29 DNA, Int. J. Mol. Sci. 11 (12) (2010) 5129–5142.
[80] E.J. Sacho, F.A. Kadyrov, P. Modrich, T.A. Kunkel, D.A. Erie, Direct visualization of asymmetric adenine-nucleotide-induced conformational changes in MutL alpha, Mol. Cell 29 (1) (2008) 112–121.
[81] J.M. Eeftens, A.J. Katan, M. Kschonsak, M. Hassler, L. de Wilde, E.M. Dief, C.H. Haering, C. Dekker, Condensin Smc2-Smc4 dimers are flexible and dynamic, Cell Rep. 14 (8) (2016) 1813–1818.
[82] M.C. Hall, H. Wang, D.A. Erie, T.A. Kunkel, High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer, J. Mol. Biol. 312 (4) (2001) 637–647.
[83] I. Tessmer, T. Moore, R.G. Lloyd, A. Wilson, D.A. Erie, S. Allen, S.J. Tendler, AFM studies on the role of the protein RdgC in bacterial DNA recombination, J. Mol. Biol. 350 (2) (2005) 254–262.
[84] K. Umemura, T. Okada, R. Kuroda, Cooperativity and intermediate structures of single-stranded DNA binding-assisted RecA-single-stranded DNA complex formation studied by atomic force microscopy, Scanning 27 (1) (2005) 35–43.
[85] S.C. Kowalczykowski, L.S. Paul, N. Lonberg, J.W. Newport, J.A. McSwiggen, P.H. von Hippel, Cooperative and noncooperative binding of protein ligands to nucleic acid lattices: experimental approaches to the determination of thermodynamic parameters, Biochemistry 25 (6) (1986) 1226–1240.
[86] F. Moreno-Herrero, L. Holtzer, D.A. Koster, S. Shuman, C. Dekker, N.H. Dekker, Atomic force microscopy shows that vaccinia topoisomerase IB generates filaments on DNA in a cooperative fashion, Nucl. Acids Res. 33 (18) (2005) 5945–5953.
[87] Y.F. Dufrene, T. Ando, R. Garcia, D. Alsteens, D. Martinez-Martin, A. Engel, C. Gerber, D.J. Muller, Imaging modes of atomic force microscopy for application in molecular and cell biology, Nat. Nanotechnol. 12 (4) (2017) 295–307.
[88] M. Guthold, M. Bezanilla, D.A. Erie, B. Jenkins, H.G. Hansma, C. Bustamante, Following the assembly of RNA polymerase-DNA complexes in aqueous solutions with the scanning force microscope, Proc. Natl. Acad. Sci. U. S. A. 91 (26) (1994) 12927–12931.
[89] C. Bustamante, M. Guthold, X. Zhu, G. Yang, Facilitated target location on DNA by individual Escherichia coli RNA polymerase molecules observed with the scanning force microscope operating in liquid, J. Biol. Chem. 274 (24) (1999) 16665–16668.
[90] M. Guthold, X. Zhu, C. Rivetti, G. Yang, N.H. Thomson, S. Kasas, H.G. Hansma, B. Smith, P.K. Hansma, C. Bustamante, Direct observation of one-dimensional diffusion and transcription by Escherichia coli RNA polymerase, Biophys. J. 77 (4) (1999) 2284–2294.
[91] S. Kasas, N.H. Thomson, B.L. Smith, H.G. Hansma, X. Zhu, M. Guthold, C. Bustamante, E.T. Kool, M. Kashlev, P.K. Hansma, Escherichia coli RNA polymerase activity observed using atomic force microscopy, Biochemistry 36 (3) (1997) 461–468.
[92] T. Ando, N. Kodera, E. Takai, D. Maruyama, K. Saito, A. Toda, A high-speed atomic force microscope for studying biological macromolecules, Proc. Natl. Acad. Sci. U. S. A. 98 (22) (2001) 12468–12472.
[93] T. Ando, High-speed atomic force microscopy coming of age, Nanotechnology 23 (6) (2012) 062001.
[94] L.S. Shlyakhtenko, A.Y. Lushnikov, Y.L. Lyubchenko, Dynamics of nucleosomes revealed by time-lapse atomic force microscopy, Biochemistry 48 (33) (2009) 7842–7848.
[95] A. Miyagi, T. Ando, Y.L. Lyubchenko, Dynamics of nucleosomes assessed with time-lapse high-speed atomic force microscopy, Biochemistry 50 (37) (2011) 7901–7908.
[96] L.S. Shlyakhtenko, A.Y. Lushnikov, A. Miyagi, M. Li, R.S. Harris, Y.L. Lyubchenko, Nanoscale structure and dynamics of ABOBEC3G complexes with single-stranded DNA, Biochemistry 51 (32) (2012) 6432–6440.
[97] Y. Pan, Z. Sun, A. Maiti, T. Kanai, H. Matsuo, M. Li, R.S. Harris, L.S. Shlyakhtenko, Y.L. Lyubchenko, Nanoscale characterization of interaction of APOBEC3G with RNA, Biochemistry 56 (10) (2017) 1473–1481.
[98] Y. Ebenstein, N. Gassman, S. Kim, S. Weiss, Combining atomic force and fluorescence microscopy for analysis of quantum-dot labeled protein-DNA complexes, J. Mol. Recognit. 22 (5) (2009) 397–402.
[99] H. Sanchez, R. Kanaar, C. Wyman, Molecular recognition of DNA-protein complexes: a straightforward method combining scanning force and fluorescence microscopy, Ultramicroscopy 110 (7) (2010) 844–851.
[100] D.N. Fronczek, C. Quammen, H. Wang, C. Kisker, R. Superfine, R. Taylor, D.A. Erie, I. Tessmer, High accuracy FIONA-AFM hybrid imaging, Ultramicroscopy 111 (5) (2011) 350–355.

[101] H. Sanchez, A. Kertokalio, S. van Rossum-Fikkert, R. Kanaar, C. Wyman, Combined optical and topographic imaging reveals different arrangements of human RAD54 with presynaptic and postsynaptic RAD51-DNA filaments, Proc. Natl. Acad. Sci. U. S. A. 110 (28) (2013) 11385–11390.
[102] H. Sanchez, M.W. Paul, M. Grosbart, S.E. van Rossum-Fikkert, J.H. Lebbink, R. Kanaar, A.B. Houtsmuller, C. Wyman, Architectural plasticity of human BRCA2-RAD51 complexes in DNA break repair, Nucl. Acids Res. (2017).
[103] D. Wu, P. Kaur, Z.M. Li, K.C. Bradford, H. Wang, D.A. Erie, Visualizing the path of DNA through proteins using DREEM imaging, Mol. Cell 61 (2) (2016) 315–323.
[104] N.L. Adkins, S.G. Swygert, P. Kaur, H. Niu, S.A. Grigoryev, P. Sung, H. Wang, C.L. Peterson, Nucleosome-like, single-stranded DNA (ssDNA)-histone octamer complexes and the implication for DNA double strand break repair, J. Biol. Chem. 292 (13) (2017) 5271–5281.
[105] D. Benarroch-Popivker, S. Pisano, A. Mendez-Bermudez, L. Lototska, P. Kaur, S. Bauwens, N. Djerbi, C.M. Latrick, V. Fraisier, B. Pei, A. Gay, E. Jaune, K. Foucher, J. Cherfils-Vicini, E. Aeby, S. Miron, A. Londono-Vallejo, J. Ye, M.H. Le Du, H. Wang, E. Gilson, M.J. Giraud-Panis, TRF2-mediated control of telomere DNA topology as a mechanism for chromosome-end protection, Mol. Cell 61 (2) (2016) 274–286.
[106] P. Kaur, D. Wu, J. Lin, P. Countryman, K.C. Bradford, D.A. Erie, R. Riehn, P.L. Opresko, H. Wang, Enhanced electrostatic force microscopy reveals higher-order DNA looping mediated by the telomeric protein TRF2, Sci. Rep. 6 (2016) 20513.
[107] E. Mikamo-Satoh, F. Yamada, A. Takagi, T. Matsumoto, T. Kawai, Electrostatic force microscopy: imaging DNA and protein polarizations one by one, Nanotechnology 20 (14) (2009) 145102.
[108] D. Alsteens, D.J. Muller, Y.F. Dufrene, Multiparametric atomic force microscopy imaging of biomolecular and cellular systems, Acc. Chem. Res. 50 (4) (2017) 924–931.
[109] D. Murugesapillai, M.J. McCauley, R. Huo, M.H. Nelson Holte, A. Stepanyants, L.J. Maher 3rd, N.E. Israeloff, M.C. Williams, DNA bridging and looping by HMO1 provides a mechanism for stabilizing nucleosome-free chromatin, Nucl. Acids Res. 42 (14) (2014) 8996–9004.
[110] J.A. North, M. Simon, M.B. Ferdinand, M.A. Shoffner, J.W. Picking, C.J. Howard, A.M. Mooney, J. van Noort, M.G. Poirier, J.J. Ottesen, Histone H3 phosphorylation near the nucleosome dyad alters chromatin structure, Nucl. Acids Res. 42 (8) (2014) 4922–4933.
[111] Z. Sun, H.Y. Tan, P.R. Bianco, Y.L. Lyubchenko, Remodeling of RecG helicase at the DNA replication fork by SSB protein, Sci. Rep. 5 (2015) 9625.
[112] V. Brazda, L. Haronikova, J.C. Liao, H. Fridrichova, E.B. Jagelska, Strong preference of BRCA1 protein to topologically constrained non-B DNA structures, BMC Mol. Biol. 17 (1) (2016) 14.
[113] K. Frykholm, R.P. Berntsson, M. Claesson, L. de Battice, R. Odegrip, P. Stenmark, F. Westerlund, DNA compaction by the bacteriophage protein Cox studied on the single DNA molecule level using nanofluidic channels, Nucl. Acids Res. 44 (15) (2016) 7219–7227.
[114] S. Santos, D. Billingsley, N. Thomson, Atomic force microscopy imaging of macromolecular complexes, Methods Mol. Biol. 950 (2013) 315–341.
[115] N.H. Thomson, S. Santos, L.A. Mitchenall, T. Stuchinskaya, J.A. Taylor, A. Maxwell, DNA G-segment bending is not the sole determinant of topology simplification by type II DNA topoisomerases, Sci. Rep. 4 (2014) 6158.
[116] A. Banerjee, W. Yang, M. Karplus, G.L. Verdine, Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA, Nature 434 (7033) (2005) 612–618.
[117] D.P. Norman, S.D. Bruner, G.L. Verdine, Coupling of substrate recognition and catalysis by a human base-excision DNA repair protein, J. Am. Chem. Soc. 123 (2) (2001) 359–360.
[118] S.D. Bruner, D.P. Norman, G.L. Verdine, Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA, Nature 403 (6772) (2000) 859–866.
[119] A. Maiti, M.T. Morgan, E. Pozharski, A.C. Drohat, Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition, Proc. Natl. Acad. Sci. U. S. A. 105 (26) (2008) 8890–8895.
[120] L. Zhang, X. Lu, J. Lu, H. Liang, Q. Dai, G.L. Xu, C. Luo, H. Jiang, C. He, Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA, Nat. Chem. Biol. 8 (4) (2012) 328–330.
[121] T. Hollis, Y. Ichikawa, T. Ellenberger, DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA, EMBO J. 19 (4) (2000) 758–766.
[122] M.R. Sawaya, R. Prasad, S.H. Wilson, J. Kraut, H. Pelletier, Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism, Biochemistry 36 (37) (1997) 11205–11215.
[123] G. Natrajan, M.H. Lamers, J.H. Enzlin, H.H. Winterwerp, A. Perrakis, T.K. Sixma, Structures of Escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates, Nucl. Acids Res. 31 (16) (2003) 4814–4821.
[124] G. Obmolova, C. Ban, P. Hsieh, W. Yang, Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA, Nature 407 (6805) (2000) 703–710.
[125] M.H. Lamers, A. Perrakis, J.H. Enzlin, H.H. Winterwerp, N. de Wind, T.K. Sixma, The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch, Nature 407 (6805) (2000) 711–717.
[126] J. van Noort, F. Orsini, A. Eker, C. Wyman, B. de Grooth, J. Greve, DNA bending by photolyase in specific and non-specific complexes studied by atomic force microscopy, Nucl. Acids Res. 27 (19) (1999) 3875–3880.
[127] A. Mees, T. Klar, P. Gnau, U. Hennecke, A.P. Eker, T. Carell, L.O. Essen, Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair, Science 306 (5702) (2004) 1789–1793.
[128] E.M. Duguid, P.A. Rice, C. He, The structure of the human AGT protein bound to DNA and its implications for damage detection, J. Mol. Biol. 350 (4) (2005) 657–666.
[129] D.H. Kim, H. Im, J.G. Jee, S.B. Jang, H.J. Yoon, A.R. Kwon, S.M. Kang, B.J. Lee, beta-Arm flexibility of HU from Staphylococcus aureus dictates the DNA-binding and recognition mechanism, Acta Crystallogr. D Biol. Crystallogr. 70 (Pt 12) (2014) 3273–3289.
[130] K.K. Swinger, K.M. Lemberg, Y. Zhang, P.A. Rice, Flexible DNA bending in HU-DNA cocrystal structures, EMBO J. 22 (14) (2003) 3749–3760.
[131] P.A. Rice, S. Yang, K. Mizuuchi, H.A. Nash, Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn, Cell 87 (7) (1996) 1295–1306.
[132] G.H. Seong, E. Kobatake, K. Miura, A. Nakazawa, M. Aizawa, Direct atomic force microscopy visualization of integration host factor-induced DNA bending structure of the promoter regulatory region on the Pseudomonas TOL plasmid, Biochem. Biophys. Res. Commun. 291 (2) (2002) 361–366.
[133] R. Sieira, D.J. Comerci, L.I. Pietrasanta, R.A. Ugalde, Integration host factor is involved in transcriptional regulation of the Brucella abortus virB operon, Mol. Microbiol. 54 (3) (2004) 808–822.
[134] A. Chakraborty, S. Lyonnais, F. Battistini, A. Hospital, G. Medici, R. Prohens, M. Orozco, J. Vilardell, M. Sola, DNA structure directs positioning of the mitochondrial genome packaging protein Abf2p, Nucl. Acids Res. 45 (2) (2017) 951–967.
[135] R. Sanchez-Giraldo, F.J. Acosta-Reyes, C.S. Malarkey, N. Saperas, M.E. Churchill, J.L. Campos, Two high-mobility group box domains act together to underwind and kink DNA, Acta Crystallogr. D Biol. Crystallogr. 71 (Pt 7) (2015) 1423–1432.
[136] M.E. Mysiak, C. Wyman, P.E. Holthuizen, P.C. van der Vliet, NFI and Oct-1 bend the Ad5 origin in the same direction leading to optimal DNA replication, Nucl. Acids Res. 32 (21) (2004) 6218–6225.
[137] J.D. Klemm, M.A. Rould, R. Aurora, W. Herr, C.O. Pabo, Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules, Cell 77 (1) (1994) 21–32.
[138] M.E. Guazzaroni, T. Krell, P. Gutierrez del Arroyo, M. Velez, M. Jimenez, G. Rivas, J.L. Ramos, The transcriptional repressor TtgV recognizes a complex operator as a tetramer and induces convex DNA bending, J. Mol. Biol. 369 (4) (2007) 927–939.
[139] D. Lu, S. Fillet, C. Meng, Y. Alguel, P. Kloppsteck, J. Bergeron, T. Krell, M.T. Gallegos, J. Ramos, X. Zhang, Crystal structure of TtgV in complex with its DNA operator reveals a general model for cooperative DNA binding of tetrameric gene regulators, Genes Dev. 24 (22) (2010) 2556–2565.
[140] D.A. Erie, G. Yang, H.C. Schultz, C. Bustamante, DNA bending by Cro protein in specific and nonspecific complexes: implications for protein site recognition and specificity, Science 266 (5190) (1994) 1562–1566.
[141] R.G. Brennan, S.L. Roderick, Y. Takeda, B.W. Matthews, Protein-DNA conformational changes in the crystal structure of a lambda Cro-operator complex, Proc. Natl. Acad. Sci. U. S. A. 87 (20) (1990) 8165–8169.
[142] R.A. Albright, B.W. Matthews, Crystal structure of lambda-Cro bound to a consensus operator at 3.0 A resolution, J. Mol. Biol. 280 (1) (1998) 137–151.
[143] R.A. Garcia, C.J. Bustamante, N.O. Reich, Sequence-specific recognition of cytosine C5 and adenine N6 DNA methyltransferases requires different deformations of DNA, Proc. Natl. Acad. Sci. U. S. A. 93 (15) (1996) 7618–7622.
[144] S. Klimasauskas, S. Kumar, R.J. Roberts, X. Cheng, HhaI methyltransferase flips its target base out of the DNA helix, Cell 76 (2) (1994) 357–369.
[145] K.C. Dong, J.M. Berger, Structural basis for gate-DNA recognition and bending by type IIA topoisomerases, Nature 450 (7173) (2007) 1201–1205.
[146] T.J. Wendorff, B.H. Schmidt, P. Heslop, C.A. Austin, J.M. Berger, The structure of DNA-bound human topoisomerase II alpha: conformational mechanisms for coordinating inter-subunit interactions with DNA cleavage, J. Mol. Biol. 424 (3–4) (2012) 109–124.

Appendix E Dancing on DNA tightropes: Watching Repair Proteins Interrogate DNA in Real Time

Mini critical review on the uses of the DNA tightrope assay; excerpt from article originally published in *Microbial Cell*. Ref. [135]: Klein H. L., Ang K., Arkin M. R., Beckwitt E. C., *et al*. Guidelines for DNA recombination and repair studies: Mechanistic assays of DNA repair processes. *Microb Cell* **6**, 65-101, doi:10.15698/mic2019.01.665 (2019).

microbial cell

Review

www.microbialcell.com

Guidelines for DNA recombination and repair studies: Mechanistic assays of DNA repair processes

Hannah L Klein[1,*], Kenny K.H. Ang[2], Michelle R. Arkin[2], Emily C. Beckwitt[3,4], Yi-Hsuan Chang[5], Jun Fan[6], Youngho Kwon[7,8], Michael J. Morten[1], Sucheta Mukherjee[9], Oliver J. Pambos[6], Hafez el Sayyed[6], Elizabeth S. Thrall[10], João P. Vieira-da-Rocha[9], Quan Wang[11], Shuang Wang[12,13], Hsin-Yi Yeh[5], Julie S. Biteen[14], Peter Chi[5,15], Wolf-Dietrich Heyer[9,16], Achillefs N. Kapanidis[6], Joseph J. Loparo[10], Terence R. Strick[12,13,17], Patrick Sung[7,8], Bennett Van Houten[3,18,19], Hengyao Niu[11,*] and Eli Rothenberg[1,*]

[1] New York University School of Medicine, Department of Biochemistry and Molecular Pharmacology, New York, NY 10016, USA.
[2] Small Molecule Discovery Center and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, USA.
[3] Program in Molecular Biophysics and Structural Biology, University of Pittsburgh, Pittsburgh, PA 15261, USA.
[4] The University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, PA 15213, USA.
[5] Institute of Biochemical Sciences, National Taiwan University, NO. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan.
[6] Biological Physics Research Group, Clarendon Laboratory, Department of Physics, University of Oxford, Oxford, OX1 3PU, UK.
[7] Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06520, USA.
[8] Department of Biochemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, Texas 78229, USA.
[9] Department of Microbiology and Molecular Genetics, University of California, Davis, CA 95616, USA.
[10] Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Boston, MA 02115, USA.
[11] Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, IN 47405, USA.
[12] Ecole Normale Supérieure, Institut de Biologie de l'Ecole Normale Supérieure (IBENS), CNRS, INSERM, PSL Research University, 75005 Paris, France.
[13] Institut Jacques Monod, CNRS, UMR7592, University Paris Diderot, Sorbonne Paris Cité F-75205 Paris, France.
[14] Departments of Chemistry and Biophysics, University of Michigan, Ann Arbor, MI 48109, USA.
[15] Institute of Biological Chemistry, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 11529, Taiwan.
[16] Department of Molecular and Cellular Biology, University of California, Davis, CA 95616, USA.
[17] Programme Equipe Labellisées, Ligue Contre le Cancer, 75013 Paris, France.
[18] Program in Molecular Biophysics and Structural Biology, University of Pittsburgh, Pittsburgh, PA 15261, USA.
[19] Department of Pharmacology and Chemical Biology, University of Pittsburgh, Pittsburgh, PA 15261, USA.

* Corresponding Authors:
Hannah Klein, New York University School of Medicine, New York, NY USA; E-mail: hannah.klein@nyumc.org;
Hengyao Niu, Indiana University, Bloomington, IN USA; E-mail: hniu@indiana.edu;
Eli Rothenberg, New York University School of Medicine, New York, NY USA; E-mail: eli.rothenberg@nyumc.org

ABSTRACT **Genomes are constantly in flux, undergoing changes due to recombination, repair and mutagenesis. *In vivo*, many of such changes are studies using reporters for specific types of changes, or through cytological studies that detect changes at the single-cell level. Single molecule assays, which are reviewed here, can detect transient intermediates and dynamics of events. Biochemical assays allow detailed investigation of the DNA and protein activities of each step in a repair, recombination or mutagenesis event. Each type of assay is a powerful tool but each comes with its particular advantages and limitations. Here the most commonly used assays are reviewed, discussed, and presented as the guidelines for future studies.**

doi: *10.15698/mic2019.01.665*
Received *originally: 16.08.2018;*
in revised form: 25.10.2018,
Accepted *31.10.2018,*
Published *07.01.2019.*

Keywords: *chromatin dynamics, chromosome rearrangements, crossovers, DNA breaks, DNA helicases, DNA repair centers, DNA repair synthesis, DNA resection, double strand break repair, DSBs, endonuclease protection assay, genome instability, gross chromosome rearrangements, fluorescent proteins, FRET, homologous recombination, mismatch repair, nonhomologous end joining, nucleotide excision repair, PALM, photoactivated fluorescent proteins, recombinase filament assembly, single-molecule, single-particle tracking, super resolution, structure-selective endonucleases, synthesis-dependent strand annealing, transcription coupled repair.*

Abbreviations:
2-AP *– 2-aminopurine,* ***BER*** *– base excision repair,* ***CMOS*** *– complementary metal–oxide–semiconductor,* ***dHJ*** *– double Holliday junction,* ***D-loop*** *– displacement loop,* ***DSB*** *– double-strand break,* ***DSBR*** *– double strand break repair,* ***dsDNA*** *– double stranded DNA,* ***EMCCD*** *– electron charge coupled device,* ***FP*** *– fluorescent protein,* ***FRAP*** *– fluorescence redistribution after photobleaching,* ***FRET*** *– Förster resonance energy transfer,* ***FROS*** *– fluorescent repressor operating system,* ***GFP*** *– green fluorescent protein,* ***HR*** *– homologous recombination,* ***HMM*** *– hidden Markov model,* ***HTS*** *– high throughput screening,* ***MMR*** *– mismatch repair,* ***MMS*** *– methyl methanesulfonate,* ***MSD*** *– mean-square displacement,* ***NER*** *– nucleotide excision repair,* ***NHEJ*** *– nonhomologous end joining,* ***OD*** *– optical density,* ***PAFPs*** *– photoactivable fluorescent proteins,* ***PALM*** *– photoactivated localization microscopy,* ***QDots*** *– quantum dots,* ***RNAP*** *– RNA polymerase,* ***SDSA*** *– synthesis-dependent strand annealing,* ***smFRET*** *– single molecule Förster resonance energy transfer,* ***ssDNA*** *– single stranded DNA,* ***SPT*** *– single-particle tracking,* ***TCR*** *– transcription coupled repair,* ***TIRF*** *– total internal reflection fluorescence,* ***TLS*** *– translesion synthesis,* ***UV*** *– ultraviolet,* ***WT*** *–wild type,* ***YFP*** *– yellow fluorescent protein.*

INTRODUCTION

Genomes are constantly subject to DNA damage arising from endogenous and exogenous sources that result in single or double stranded breaks, modified bases, and chromatin changes, among others. To protect the genome, cells have an arsenal of repair mechanisms to sue, the specific mechanism dependent on the type of damage and its context. Our understanding of the myriad repair pathways has come from genetic studies to identify genes encoding proteins for DNA repair and the consequences of loss of these functions, *in vivo* genetic and physical assays to determine the consequences of failure to repair, cytological assays to interrogate protein interactions and real time events, and *in vitro* biochemical assays to determine the substrate and repair events, and the molecular intermediates in repair.

In a separate guideline article, we have reviewed genetic, molecular and cytological assays for repair. In this guideline article mechanistic assays are presented, specifically single molecule assays and biochemical assays. Single molecule assays can be applied to *in vivo* or *in vitro* situations. Single molecule fluorescence and PALM (photoactivated localization microscopy) imaging are used to study the position and dynamics of tagged proteins interacting with DNA substrates that are induced by external stimuli. Movement of proteins on DNA molecules, using DNA tightropes or DNA nanomanipulation and a magnetic trap allows visualization of DNA topology changes resulting from protein interaction with the DNA molecules. Both types of approaches have led to a detailed understanding of repair processes and in some cases have challenged the current models of repair.

Biochemical assays permit detailed investigation of DNA protein interactions. Reactions mimicking the proposed intermediates in homologous recombination (HR) are the focus of the guidelines here. From the initial step in recombination, the assembly of the presynaptic filament to the formation of the D-loop, followed by extension of the D-loop from the primer terminus, these reactions are studied *in vitro* using substrates and purified proteins. The pro-

posed intermediates are often derived from *in vivo* genetic experiments and tested *in vitro*. The *in vitro* results then inform further *in vivo* biological experiments. HR involves DNA helicases and nucleases. Assays for helicases are included here, which represented key steps in the HR process. Finally, structure-selective endonucleases are needed at several steps in the HR process. Here, different types of substrates and assays for joint molecule resolution are presented.

These guidelines should be useful for the application of these approaches to many areas of DNA repair. Individual author contributions and contact information are available in Supplementary Table 1.

SINGLE MOLECULE ASSAYS FOCUSING ON DNA REPAIR

Single molecule assays are powerful tools that can be used to investigate the activity of proteins on DNA. They bypass the need to synchronize initiation events and enable the detection of transient intermediates that are otherwise lost to ensemble averaging. This section describes several single molecule techniques and some of the insights into DNA repair that have been directly made from the minute level of detail that these assays are able to provide (**Box1**).

Dancing on DNA tightropes: watching repair proteins interrogate DNA in real time

In order to understand how DNA repair proteins find damaged sites in a vast excess of non-damaged DNA, the field of DNA repair has moved to various single molecule approaches allowing direct visualization of proteins interacting with their DNA substrates [1]. These single molecule techniques can provide unique insights into population trends without losing detailed information on individual particles or events [2]. An optical platform consisting of DNA tightropes was developed by Neil Kad at the University of Vermont and first used to study bacterial nucleotide excision repair (NER) proteins [3, 4] and base excision repair (BER) glycosylases [5]. This DNA tightrope assay takes a similar approach to the DNA curtain setup developed by Dr. Eric Greene and colleagues [6, 7] with one important difference. The tightrope itself is established by suspending long molecules of double stranded (ds) DNA (~90% contour length) between poly-L-lysine coated five micron beads dispersed in a flow cell (**Figure 1A**). Visualizing repair proteins of interest up off the surface requires labels with bright fluorescent signals, and real-time imaging requires photostability over long periods. To accomplish these two needs, repair proteins are conjugated to quantum dots (Qdots) with appropriate antibodies (**Figure 1B**) and added to the flow cell. Interactions are recorded in real time, in the absence of flow, using oblique angle fluorescence on a total internal reflection fluorescence (TIRF) microscope with a CMOS or EMCCD camera (**Figure 1C, D**) [8]. Here, we will discuss the advantages and limitations of the DNA tightrope assay, current applications, and potential new directions.

The tightrope assay has its own exceptional strengths. Bringing the DNA up off the bottom of the flow cell overcomes surface interactions that can arise from DNA being in contact with a phospholipid layer, as well as it assures the observer that the Qdots being monitored are attached to DNA repair proteins engaged with the DNA and not proteins or Qdots simply sticking to the surface. Because the DNA is suspended on both ends, once the proteins of interest are added, they can be observed in the absence of flow. Finally, this optical platform allows for the use of long DNA substrates and the potential to engineer multiple site-specific lesions that can be marked with Qdots [8]. The use of Qdots, however, also presents some potential challenges.

Relatively bulky labeling strategies using large Qdots and antibodies (**Figure 1B**) may sterically hinder protein interactions with DNA and/or other proteins. Despite this potential problem, we have been able to observe three-

BOX 1: SINGLE MOLECULE ASSAYS FOCUSING ON DNA REPAIR

DNA tightropes to watch repair proteins interrogate DNA | The method of DNA tightropes to directly visualize proteins interacting with DNA substrates is described. Advantages of this method are presented with examples of target searches by DNA repair proteins.

Single-molecule (Förster resonance energy transfer) FRET illuminates the non-homologous end joining process in vitro | smFRET is used to study the details of NHEJ and deduce causes of aberrant end joining.

Single molecule imaging to study mismatch repair in living cells | Live cell single-molecule fluorescence is used to study MutS in bacterial cells. The positioning and dynamics of proteins can be assessed and responses to external stimuli determined to understand a repair process at the nanometer scale.

Single molecule DNA nanomanipulation | Use of a magnetic trap to observe real-time changes in DNA topology and structure from protein interactions. Here it is used to study MutS in bacteria.

Single molecule PALM imaging | A description of PALM and its application to translesion polymerases in living bacterial cells is presented.

Tracking-PALM direct single-molecule imaging | Combining single-molecule tracking with PALM has led to a localization-based super-resolution imaging method. Here use of this method to study DNA repair in living bacteria is presented.

color Qdot-labeled NER UvrABC complexes moving together on DNA [9]. Controls of non-conjugated Qdots and optimization of protein:antibody:Qdot ratios are required for such experiments. The size of the Qdot (~10-15 nm) and inherent rotational drag must also be considered when analyzing the diffusive behaviors of proteins on DNA. The use of oblique angle illumination enhances signal-to-noise over epifluorescence microscopy and resolution can be further improved by fitting Gaussians to the intensity profiles along a kymograph (**Figure 1E**). Movies can be collected as fast as 100 frames per second and the mean positional accuracy for a Qdot-labeled protein has been reported as 6 ± 3 nm [10].

The DNA tightrope assay can be used to answer several major questions about protein-DNA interactions. First, and perhaps most importantly for this method, is the question of modes of target search (**Figure 1F**) [11]. Resolution limits prevent observation of very short-range motion below 100-200 bp, but motion above this scale can be investigated in depth. Movies of protein-DNA interactions are converted to kymographs and subsequent mathematical analyses of observed linear diffusive behavior can provide insight into the molecular basis for these interactions. Mean square displacement analysis of particle motion is used to calculate the diffusion coefficient *D* and anomalous diffusion exponent α, providing information about rates and nature of the diffusion process (**Figure 1E**). Surprisingly we have found that several repair proteins, including Rad4 [12] and PARP1 [13], undergo anomalous diffusion, showing highly constrained motion around the site of damage. In addition, Dr. Susan Wallace's group has shown that aromatic side chains of BER glycosylases caused pausing at damaged sites in DNA [5, 14]. Furthermore, the cohesion protein SA1 was observed to alternate between fast and slow diffusion and this was dependent on telomeric sequences used in the DNA tightropes [15].

The use of orthogonal labeling strategies (i.e. Qdots with distinct emission spectra and conjugation schemes, **Figure 1B**) can be used to answer questions about colocalization and other interactions on DNA. Dimerization or interaction of two (or more) DNA repair proteins can be observed by separate and different labeling of the proteins of interest. Furthermore, such experiments can detect changes in dynamic behavior of proteins in the absence/presence of other DNA repair partners. For example, the eukaryotic NER recognition protein UV-DDB was observed to dimerize on UV-damaged DNA and abasic DNA [10]. In another example, UvrB was only observed binding to DNA tightropes in complexes containing UvrC or UvrA [9, 16]. To determine if proteins colocalize with target lesions, site specific arrays of DNA damage can be engineered with a biotinylated base proximal to the lesion and labeled with a streptavidin-coated Qdot orthologous from the labeled proteins [8]. In this way, UV-DDB [10] and PARP1 [13] were observed colocalizing with abasic sites along DNA tightropes. However, limits in spatial resolution dictate that direct interactions should be confirmed with complementary methods.

The DNA tightrope assay has made important contributions to the study of DNA repair proteins from both microbial systems and more complex multicellular organisms. Use of this optical platform will continue to foster progress in the field as the method is improved and modified to suit newer needs. For example, assembly of nucleosomes along DNA can be used to study chromatin [17]. Furthermore, incubation of DNA tightropes with nuclear extracts will allow for the study of specifically labeled proteins in the context of all their interacting partners [18]. The future holds great promise as single molecule detection of DNA repair proteins dancing on DNA occurs in even more physiologically relevant settings, and even within living *Escherichia coli* cells [19].

Single-molecule (Förster resonance energy transfer) FRET illuminates the non-homologous end joining process *in vitro*

Overview

The central premise behind single-molecule experiments is to avoid losing information through ensemble averaging. DNA:protein interactions are well suited to be studied at a single-molecule resolution, in part, due to the relatively facile isolation and detection of individual DNA molecules. Chromosomal double strand breaks (DSBs) are arguably the most cytotoxic form of DNA damage, and are fatal to a cell if left unchecked. Non-homologous end joining (NHEJ) dominates over HR during G1 in mammalian cells, most notably due to the lack of a sister chromatid template to complete HR, but it is known to generate errors that are also extremely damaging to the cell [20, 21]. DSBs can produce DNA ends with varied chemistries, and the NHEJ machinery includes end processing enzymes to efficiently join different types of broken ends together [22-24]. However, there are certain DNA substrates that are more prone to incorrectly repair DSBs, and the reasons behind this are still unclear. Single-molecule Förster resonance energy transfer (smFRET) experiments using total internal reflection fluorescence microscopy (TIRFm) are ideally equipped to accurately quantify rate constants and identify transient intermediates that are otherwise hidden in an ensemble. smFRET is therefore well suited to illuminate the subtleties of the NHEJ mechanism and deduce the causes of aberrant end joining.

Description of method/assay

To study NHEJ using smFRET, fluorescently labeled DNA substrates can be immobilized to a surface, and the intensity of the fluorescent dyes can be recorded throughout the end joining process [21]. The two pieces of DNA are labeled with two different fluorophores, Cy3 and Cy5, which act as an energy donor and acceptor respectively. FRET is only likely to occur when these dyes are close to each other, therefore the FRET response can be interpreted in terms of the relative distance between two linear DNA molecules, which are analogous to the ends produced by a DSB.

A sample chamber is created between a coverslip and glass slide, and the internal walls are passivated by a PEG

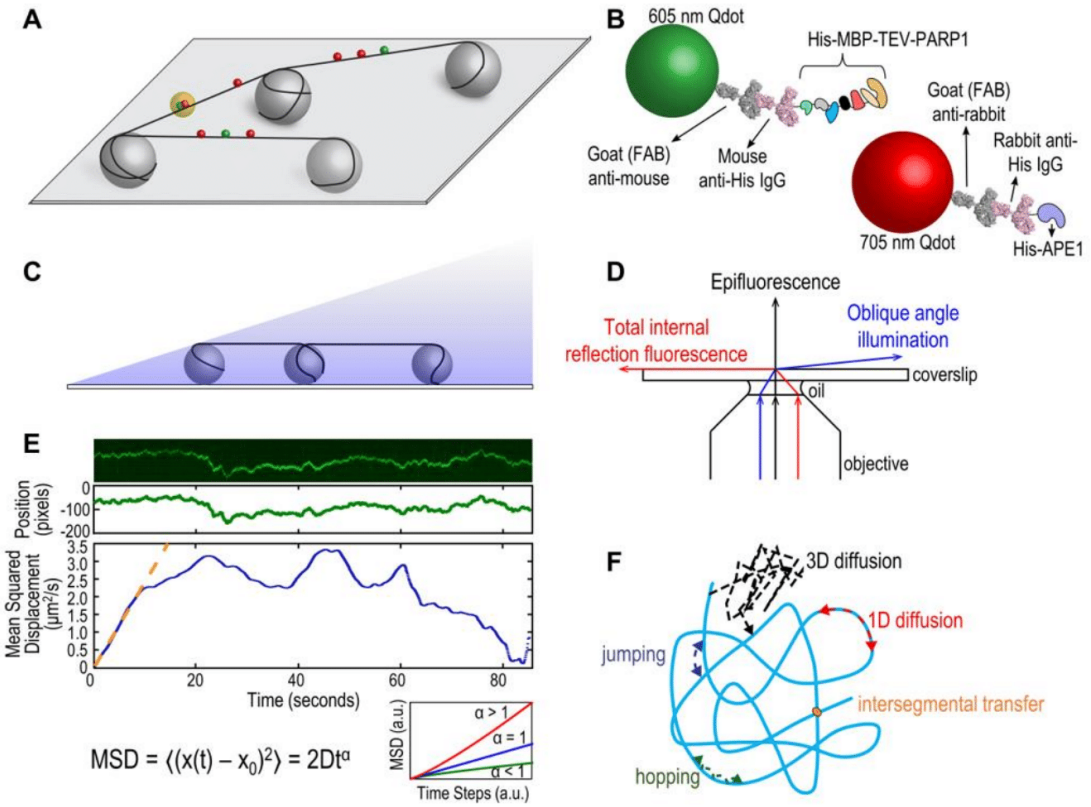

FIGURE 1: (A) Schematic of DNA tightrope setup. Long DNA molecules are suspended between 5 µm poly-L-lysine-coated silica beads on a glass coverslip. Qdot-labeled proteins bound to DNA shown in red and green (see B); colocalized particles highlighted in yellow. Adapted with permission from [13]. **(B) Two orthogonal Qdot-protein labeling strategies.** Top: A 605 nm Qdot (green) with conjugated anti-mouse secondary antibody (grey) bound to a mouse anti-His primary antibody (pink), bound to a His-tagged protein. Bottom: A 705 nm Qdot (red) with conjugated anti-rabbit secondary antibody (grey) bound to a rabbit anti-His primary antibody (pink), bound to a His-tagged protein. Adapted with permission from [13]. **(C) DNA tightropes in a flow cell with oblique angle illumination.** Adapted with permission from [8]. **(D) Ray diagram showing incident laser light paths for epifluorescence (black), TIRF at the critical angle (red), and oblique angle illumination (blue).** Adapted with permission from [8]. **(E) Top:** Sample kymograph of a Qdot-labeled protein displaying random linear diffusion on a DNA tightrope. Y axis, position; X axis, time. **Middle:** 1D Gaussian fittings of the light intensity profile at each time point of the above kymograph shown as position in pixels vs. time. 1 pixel = 46 nm. **Bottom:** Mean squared displacement (MSD) vs. time. The initial linear portion of the MSD plot is fit to the equation MSD = $2Dt^{\alpha}$ (orange line). Inset: Sub-types of 1D diffusion defined by α values. Super-diffusion (red), random diffusion (blue), sub-diffusion (green). Adapted with permission from [11]. **(F) Modes of protein-DNA interaction.** Search strategies typically involve some combination of: 3D diffusion in solution (black), 1D linear diffusion (red), jumping (blue) or hopping (green) between DNA segments, and intersegmental transfer (orange). Adapted with permission from [11].

surface to minimize non-specific binding. DNA is covalently bound to a biotin moiety which interacts with neutravidin molecules on the modified surface of the glass coverslip [25]. Single-molecule resolution is achieved by only sparsely populating the slide with an immobilized DNA substrate so that each pixel corresponds to a region on the slide with only one fluorophore [26]. Typically, an incubation of low picomolar concentrations of the biotinylated DNA is sufficient to produce a surface that is populated by many, but distinct, DNA molecules. The number of immobilized DNA molecules in each pixel on screen can be confirmed by photobleaching experiments to show that the majority of high intensity spots measure the emission from a single dye only. A second DNA structure can then be introduced to the sample chamber, along with the necessary proteins to carry out the end joining process.

The initial joining of two DNA ends by NHEJ proteins form a paired end complex (PEC) as shown in **Figure 2A**, and can be monitored in a number of ways: the number of FRET pairs observed can be used to quantify the yield of the end joining reaction; the changes in FRET efficiency during PEC formation allows the movement of the DNA ends to be observed; and the measurement of the dwell times in between these movements can infer the stability

SUMMARY

The maintenance of genomic information and genomic integrity lies at the core of all organismal propagation, development and survival. The past two decades have witnessed the emergence of several new and powerful physical approaches collectively termed single-molecule techniques. Utilizing these methods for studying biological systems provides many new features that are otherwise masked due to averaging in ensemble measurements, thus providing previously unattainable data and new mechanistic insights. Herein we provide a concise description and procedures on the use of several next-generation single-molecule techniques, assays and tools that are used to study key molecular mechanisms and pathways in DNA repair, and address fundamental questions in the field. These include methods such as super-resolution localization microscopy for real-time tracking of individual molecules in live cells, single-molecule tracking in vitro assays, single-molecule manipulation, and single-molecule FRET. Beyond their practical description, we sought to highlight both the strengths and limitations of each technique to give in context explanations of how each method should be employed to investigate DNA repair mechanisms.

ACKNOWLEDGMENTS

HLK was supported by NIH grant R01-CA146940. JB acknowledges support from NIH grant R21-GM-128022-01. PC was supported by the Taiwan Ministry of Science and Technology grant MOST 105-2314-B-002-073-MY4. W-DH was supported by NIH grants R01-GM58015, R01-CA92276, DOD grant W81XWH-14-1-0435 and T32-GM099608 (SM). BVH was supported by NIH grants R01-ES019566 (BVH) and T32-GM088119 (ECB). ANK was supported by grants to ANK from the European Research Council (261227), the Wellcome Trust (110164/Z/15/Z), the UK BBSRC (BB/H01795X/1, BB/J00054X/1, and BB/N018656/1). JJL was funded by National Institutes of Health grants R01 GM114065 (to JJL) and F32 GM113516 (to EST). TRS would like to acknowledge grants RepOne and PrTxConf (French ANR), NanoRep (Université PSL) as well as the "Equipe Labelisée" program of the French League Against Cancer. PS was supported by NIH grant R01-ES00761. HN was supported by NIH grant R35-GM124765. ER is funded by grants from NIH: R01 GM108119, ACS: 130304-RSG-16-241-01-DMC and V Foundation for Cancer Research: D2018-020.

SUPPLEMENTAL MATERIAL

All supplemental data for this article are available online at www.microbialcell.com.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Please cite this article as: Hannah L Klein, Kenny K.H. Ang, Michelle R. Arkin, Emily C. Beckwitt, Yi-Hsuan Chang, Jun Fan, Youngho Kwon, Michael J. Morten, Sucheta Mukherjee, Oliver J. Pambos, Hafez el Sayyed, Elizabeth S. Thrall, João P. Vieira-da-Rocha, Quan Wang, Shuang Wang, Hsin-Yi Yeh, Julie S. Biteen, Peter Chi, Wolf-Dietrich Heyer, Achillefs N. Kapanidis, Joseph J. Loparo, Terence R. Strick, Patrick Sung, Bennett Van Houten, Hengyao Niu and Eli Rothenberg (**2019**). Guidelines for DNA recombination and repair studies: Mechanistic assays of DNA repair processes. **Microbial Cell** 6(1): 65-101. doi: 10.15698/mic2019.01.665

REFERENCES

1. Kad NM, Van Houten B, editors (**2014**). Single molecule approaches: watching DNA repair one molecule at a time. **DNA Repair** 20:1. doi: 10.1016/j.dnarep.2014.06.007

2. Monachino E, Spenkelink LM, van Oijen AM (**2017**). Watching cellular machinery in action, one molecule at a time. **J Cell Biol** 216(1): 41-51. doi: 10.1083/jcb.201610025

3. Kad NM, Wang H, Kennedy GG, Warshaw DM, Van Houten B (**2010**). Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by single- molecule imaging of quantum-dot-labeled proteins. **Mol Cell** 37(5): 702-713. doi: 10.1016/j.molcel.2010.02.003

4. Van Houten B, Kad N (**2014**). Investigation of bacterial nucleotide excision repair using single-molecule techniques. **DNA Repair** 20: 41-48. doi: 10.1016/j.dnarep.2013.10.012

5. Dunn AR, Kad NM, Nelson SR, Warshaw DM, Wallace SS (**2011**). Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. **Nucleic Acids Res** 39(17): 7487-7498. doi: 10.1093/nar/gkr459

6. Graneli A, Yeykal CC, Robertson RB, Greene EC (**2006**). Long-distance lateral diffusion of human Rad51 on double-stranded DNA. **Proc Natl Acad Sci U S A** 103(5): 1221-1226. doi: 10.1073/pnas.0508366103

7. Gorman J, Chowdhury A, Surtees JA, Shimada J, Reichman DR, Alani E, Greene EC (**2007**). Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. **Mol Cell** 28(3): 359-370. doi: 10.1016/j.molcel.2007.09.008

8. Kong M, Beckwitt EC, Springall L, Kad NM, Van Houten B (**2017**). Single-Molecule Methods for Nucleotide Excision Repair: Building a System to Watch Repair in Real Time. **Methods Enzymol** 592: 213-257. doi: 10.1016/bs.mie.2017.03.027

9. Springall L, Hughes CD, Simons M, Azinas S, Van Houten B, Kad NM (**2018**). Recruitment of UvrBC complexes to UV-induced damage in the absence of UvrA increases cell survival. **Nucleic Acids Res** 46(3): 1256-1265. doi: 10.1093/nar/gkx1244

10. Ghodke H, Wang H, Hsieh CL, Woldemeskel S, Watkins SC, Rapic-Otrin V, Van Houten B (**2014**). Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates. **Proc Natl Acad Sci U S A** 111(18): E1862-1871. doi: 10.1073/pnas.1323856111

11. Kong M, Van Houten B (**2017**). Rad4 recognition-at-a-distance: Physical basis of conformation-specific anomalous diffusion of DNA

repair proteins. **Prog Biophys Mol Biol** 127: 93-104. doi: 10.1016/j.pbiomolbio.2016.12.004

12. Kong M, Liu L, Chen X, Driscoll KI, Mao P, Bohm S, Kad NM, Watkins SC, Bernstein KA, Wyrick JJ, Min JH, Van Houten B (**2016**). Single-Molecule Imaging Reveals that Rad4 Employs a Dynamic DNA Damage Recognition Process. **Mol Cell** 64(2): 376-387. doi: 10.1016/j.molcel.2016.09.005

13. Liu L, Kong M, Gassman NR, Freudenthal BD, Prasad R, Zhen S, Watkins SC, Wilson SH, Van Houten B (**2017**). PARP1 changes from three-dimensional DNA damage searching to one-dimensional diffusion after auto-PARylation or in the presence of APE1. **Nucleic Acids Res** 45(22): 12834-12847. doi: 10.1093/nar/gkx1047

14. Nelson SR, Dunn AR, Kathe SD, Warshaw DM, Wallace SS (**2014**). Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. **Proc Natl Acad Sci U S A** 111(20): E2091-2099. doi: 10.1073/pnas.1400386111

15. Lin J, Countryman P, Chen H, Pan H, Fan Y, Jiang Y, Kaur P, Miao W, Gurgel G, You C, Piehler J, Kad NM, Riehn R, Opresko PL, Smith S, Tao YJ, Wang H (**2016**). Functional interplay between SA1 and TRF1 in telomeric DNA binding and DNA-DNA pairing. **Nucleic Acids Res** 44(13): 6363-6376. doi: 10.1093/nar/gkw518

16. Hughes CD, Wang H, Ghodke H, Simons M, Towheed A, Peng Y, Van Houten B, Kad NM (**2013**). Real-time single-molecule imaging reveals a direct interaction between UvrC and UvrB on DNA tightropes. **Nucleic Acids Res** 41(9): 4901-4912. doi: 10.1093/nar/gkt177

17. Finkelstein IJ, Visnapuu ML, Greene EC (**2010**). Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. **Nature** 468(7326): 983-987. doi: 10.1038/nature09561

18. Duzdevich D, Warner MD, Ticau S, Ivica NA, Bell SP, Greene EC (**2015**). The dynamics of eukaryotic replication initiation: origin specificity, licensing, and firing at the single-molecule level. **Mol Cell** 58(3): 483-494. doi: 10.1016/j.molcel.2015.03.017

19. Uphoff S, Reyes-Lamothe R, Garza de Leon F, Sherratt DJ, Kapanidis AN (**2013**). Single-molecule DNA repair in live bacteria. **Proc Natl Acad Sci U S A** 110(20): 8063-8068. doi: 10.1073/pnas.1301804110

20. Reid DA, Rothenberg E (**2015**). Repair of chromosomal breaks by NHEJ. **Oncotarget** 6(18): 15730-15731. doi: 10.18632/oncotarget.4593

21. Reid DA, Keegan S, Leo-Macias A, Watanabe G, Strande NT, Chang HH, Oksuz BA, Fenyo D, Lieber MR, Ramsden DA, Rothenberg E (**2015**). Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. **Proc Natl Acad Sci U S A** 112(20): E2575-2584. doi: 10.1073/pnas.1420115112

22. Conlin MP, Reid DA, Small GW, Chang HH, Watanabe G, Lieber MR, Ramsden DA, Rothenberg E (**2017**). DNA Ligase IV Guides End-Processing Choice during Nonhomologous End Joining. **Cell Rep** 20(12): 2810-2819. doi: 10.1016/j.celrep.2017.08.091

23. Chang HH, Watanabe G, Gerodimos CA, Ochi T, Blundell TL, Jackson SP, Lieber MR (**2016**). Different DNA End Configurations Dictate Which NHEJ Components Are Most Important for Joining Efficiency. **J Biol Chem** 291(47): 24377-24389. doi: 10.1074/jbc.M116.752329

24. Gu J, Lu H, Tippin B, Shimazaki N, Goodman MF, Lieber MR (**2007**). XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. **EMBO J** 26(4): 1010-1023. doi: 10.1038/sj.emboj.7601559

25. Joo C, Ha T (**2012**). Preparing sample chambers for single-molecule FRET. **Cold Spring Harb Protoc** 2012(10): 1104-1108. doi: 10.1101/pdb.prot071530

26. Roy R, Hohng S, Ha T (**2008**). A practical guide to single-molecule FRET. **Nat Methods** 5(6): 507-516. doi: 10.1038/nmeth.1208

27. Reid DA, Conlin MP, Yin Y, Chang HH, Watanabe G, Lieber MR, Ramsden DA, Rothenberg E (**2017**). Bridging of double-stranded breaks by the nonhomologous end-joining ligation complex is modulated by DNA end chemistry. **Nucleic Acids Res** 45(4): 1872-1878. doi: 10.1093/nar/gkw1221

28. Axelrod D (**2001**). Selective imaging of surface fluorescence with very high aperture microscope objectives. **J Biomed Opt** 6(1): 6-13. doi: Doi 10.1117/1.1335689

29. Aitken CE, Marshall RA, Puglisi JD (**2008**). An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. **Biophys J** 94(5): 1826-1835. doi: 10.1529/biophysj.107.117689

30. Rasnik I, McKinney SA, Ha T (**2006**). Nonblinking and long-lasting single-molecule fluorescence imaging. **Nat Methods** 3(11): 891-893. doi: 10.1038/nmeth934

31. McKinney SA, Joo C, Ha T (**2006**). Analysis of single-molecule FRET trajectories using hidden Markov modeling. **Biophys J** 91(5): 1941-1951. doi: 10.1529/biophysj.106.082487

32. Lieber MR (**2010**). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. **Annu Rev Biochem** 79(1): 181-211. doi: 10.1146/annurev.biochem.052308.093131

33. Heyes CD, Groll J, Moller M, Nienhaus GU (**2007**). Synthesis, patterning and applications of star-shaped poly(ethylene glycol) biofunctionalized surfaces. **Mol Biosyst** 3(6): 419-430. doi: 10.1039/b700055n

34. Ha T, Tinnefeld P (**2012**). Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. **Annu Rev Phys Chem** 63: 595-617. doi: 10.1146/annurev-physchem-032210-103340

35. Cordes T, Vogelsang J, Tinnefeld P (**2009**). On the mechanism of Trolox as antiblinking and antibleaching reagent. **J Am Chem Soc** 131(14): 5018-5019. doi: 10.1021/ja809117z

36. Saurabh S, Maji S, Bruchez MP (**2012**). Evaluation of sCMOS cameras for detection and localization of single Cy5 molecules. **Opt Express** 20(7): 7338-7349. doi: 10.1364/OE.20.007338

37. Di Fiori N, Meller A (**2010**). The Effect of dye-dye interactions on the spatial resolution of single-molecule FRET measurements in nucleic acids. **Biophys J** 98(10): 2265-2272. doi: 10.1016/j.bpj.2010.02.008

38. Modrich P (**2006**). Mechanisms in Eukaryotic Mismatch Repair. **J Biol Chem** 281(41): 30305-30309. doi: 10.1074/jbc.r600022200

39. Jiricny J (**2013**). Postreplicative Mismatch Repair. **Cold Spring Harb Perspect Biol** 5(4): a012633-a012633. doi: 10.1101/cshperspect.a012633

40. Li Y, Schroeder JW, Simmons LA, Biteen JS (**2018**). Visualizing bacterial DNA replication and repair with molecular resolution. **Curr Opin Microbiol** 43: 38-45. doi: 10.1016/j.mib.2017.11.009

41. Liao Y, Schroeder JW, Gao B, Simmons LA, Biteen JS (**2015**). Single-molecule motions and interactions in live cells reveal target search dynamics in mismatch repair. **Proc Natl Acad Sci U S A** 112(50): E6898-E6906. doi: 10.1073/pnas.1507386112

42. Tuson HH, Biteen JS (**2015**). Unveiling the inner workings of live bacteria using super-resolution microscopy. **Anal Chem** 87(1): 42-63. doi: 10.1021/ac5041346

43. Manley S, Gillette JM, Patterson GH, Shroff H, Hess HF, Betzig E, Lippincott-Schwartz J (**2008**). High-density mapping of single-molecule trajectories with photoactivated localization microscopy. **Nat Methods** 5(2): 155-157. doi: 10.1038/nmeth.1176

44. Subach FV, Patterson GH, Manley S, Gillette JM, Lippincott-Schwartz J, Verkhusha VV (**2009**). Photoactivatable mCherry for high-

resolution two-color fluorescence microscopy. **Nat Methods** 6(2): 153-159. doi: 10.1038/nmeth.1298

45. Wang S, Moffitt JR, Dempsey GT, Xie XS, Zhuang X (**2014**). Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging. **Proc Natl Acad Sci U S A** 111(23): 8452-8457. doi: 10.1073/pnas.1406593111

46. Jeong C, Cho W-K, Song K-M, Cook C, Yoon T-Y, Ban C, Fishel R, Lee J-B (**2011**). MutS switches between two fundamentally distinct clamps during mismatch repair. **Nat Struct Mol Biol** 18(3): 379-385. doi: 10.1038/nsmb.2009

47. Lewis JP (**1995**). Fast Normalized Cross-Correlation. In: Society CIPaPR, editor Vision Interface. pp 120–123.

48. Walsh BW, Bolz SA, Wessel SR, Schroeder JW, Keck JL, Simmons LA (**2014**). RecD2 helicase limits replication fork stress in Bacillus subtilis. **J Bacteriol** 196(7): 1359-1368. doi: 10.1128/JB.01475-13

49. Liao Y, Li Y, Schroeder JW, Simmons LA, Biteen JS (**2016**). Single-Molecule DNA Polymerase Dynamics at a Bacterial Replisome in Live Cells. **Biophys J** 111(12): 2562-2569. doi: 10.1016/j.bpj.2016.11.006

50. Strick TR, Allemand JF, Bensimon D, Bensimon A, Croquette V (**1996**). The elasticity of a single supercoiled DNA molecule. **Science** 271(5257): 1835-1837. doi: 10.1126/science.271.5257.1835

51. Strick TR, Croquette V, Bensimon D (**2000**). Single-molecule analysis of DNA uncoiling by a type II topoisomerase. **Nature** 404(6780): 901-904. doi: 10.1038/35009144

52. Revyakin A, Ebright RH, Strick TR (**2004**). Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. **Proc Natl Acad Sci U S A** 101(14): 4776-4780. doi: 10.1073/pnas.0307241101

53. Strick TR, Kawaguchi T, Hirano T (**2004**). Real-time detection of single-molecule DNA compaction by condensin I. **Curr Biol** 14(10): 874-880. doi: 10.1016/j.cub.2004.04.038

54. Strick TR, Charvin G, Dekker NH, Allemand JF, Bensimon D, Croquette V (**2002**). Tracking enzymatic steps of DNA topoisomerases using single-molecule micromanipulation. **Comptes Rendus Physique** 3(5): 595-618. doi: Pii S1631-0705(02)01347-6/Fla

doi 10.1016/S1631-0705(02)01347-6

55. Charvin G, Strick TR, Bensimon D, Croquette V (**2005**). Topoisomerase IV bends and overtwists DNA upon binding. **Biophys J** 89(1): 384-392. doi: 10.1529/biophysj.105.060202

56. Revyakin A, Liu C, Ebright RH, Strick TR (**2006**). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. **Science** 314(5802): 1139-1143. doi: 10.1126/science.1131398

57. Lerner E, Chung S, Allen BL, Wang S, Lee J, Lu SW, Grimaud LW, Ingargiola A, Michalet X, Alhadid Y, Borukhov S, Strick TR, Taatjes DJ, Weiss S (**2016**). Backtracked and paused transcription initiation intermediate of Escherichia coli RNA polymerase. **Proc Natl Acad Sci U S A** 113(43): E6562-E6571. doi: 10.1073/pnas.1605038113

58. Howan K, Smith AJ, Westblade LF, Joly N, Grange W, Zorman S, Darst SA, Savery NJ, Strick TR (**2012**). Initiation of transcription-coupled repair characterized at single-molecule resolution. **Nature** 490(7420): 431-434. doi: 10.1038/nature11430

59. Fan J, Leroux-Coyau M, Savery NJ, Strick TR (**2016**). Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. **Nature** 536(7615): 234-237. doi: 10.1038/nature19080

60. Graves ET, Duboc C, Fan J, Stransky F, Leroux-Coyau M, Strick TR (**2015**). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. **Nat Struct Mol Biol** 22(6): 452-457. doi: 10.1038/nsmb.3019

61. Strick TR, Allemand JF, Bensimon D, Croquette V (**1998**). Behavior of supercoiled DNA. **Biophys J** 74(4): 2016-2028. doi: 10.1016/S0006-3495(98)77908-1

62. Thrall ES, Kath JE, Chang S, Loparo JJ (**2017**). Single-molecule imaging reveals multiple pathways for the recruitment of translesion polymerases after DNA damage. **Nat Commun** 8(1): 2170. doi: 10.1038/s41467-017-02333-2

63. Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, Davidson MW, Lippincott-Schwartz J, Hess HF (**2006**). Imaging intracellular fluorescent proteins at nanometer resolution. **Science** 313(5793): 1642-1645. doi: 10.1126/science.1127344

64. Uphoff S, Sherratt DJ, Kapanidis AN (**2014**). Visualizing protein-DNA interactions in live bacterial cells using photoactivated single-molecule tracking. **J Vis Exp** 85. doi: 10.3791/51177

65. Datsenko KA, Wanner BL (**2000**). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. **Proc Natl Acad Sci U S A** 97(12): 6640-6645. doi: 10.1073/pnas.120163297

66. Tiruvadi Krishnan S, Moolman MC, van Laar T, Meyer AS, Dekker NH (**2015**). Essential validation methods for E. coli strains created by chromosome engineering. **J Biol Eng** 9(1): 11. doi: 10.1186/s13036-015-0008-x

67. Tokunaga M, Imamoto N, Sakata-Sogawa K (**2008**). Highly inclined thin illumination enables clear single-molecule imaging in cells. **Nat Methods** 5(2): 159-161. doi: 10.1038/nmeth1171

68. Bakshi S, Bratton BP, Weisshaar JC (**2011**). Subdiffraction-limit study of Kaede diffusion and spatial distribution in live Escherichia coli. **Biophys J** 101(10): 2535-2544. doi: 10.1016/j.bpj.2011.10.013

69. Mohapatra S, Choi H, Ge X, Sanyal S, Weisshaar JC (**2017**). Spatial Distribution and Ribosome-Binding Dynamics of EF-P in Live Escherichia coli. **MBio** 8(3): e00300-00317. doi: 10.1128/mBio.00300-17

70. Paintdakhi A, Parry B, Campos M, Irnov I, Elf J, Surovtsev I, Jacobs-Wagner C (**2016**). Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. **Mol Microbiol** 99(4): 767-777. doi: 10.1111/mmi.13264

71. Sliusarenko O, Heinritz J, Emonet T, Jacobs-Wagner C (**2011**). High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. **Mol Microbiol** 80(3): 612-627. doi: 10.1111/j.1365-2958.2011.07579.x

72. Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL, Danuser G (**2008**). Robust single-particle tracking in live-cell time-lapse sequences. **Nat Methods** 5(8): 695-702. doi: 10.1038/nmeth.1237

73. Aguet F, Antonescu CN, Mettlen M, Schmid SL, Danuser G (**2013**). Advances in analysis of low signal-to-noise images link dynamin and AP2 to the functions of an endocytic checkpoint. **Dev Cell** 26(3): 279-291. doi: 10.1016/j.devcel.2013.06.019

74. Kamentsky L, Jones TR, Fraser A, Bray MA, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW, Carpenter AE (**2011**). Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. **Bioinformatics** 27(8): 1179-1180. doi: 10.1093/bioinformatics/btr095

75. Caldas VE, Punter CM, Ghodke H, Robinson A, van Oijen AM (**2015**). iSBatch: a batch-processing platform for data analysis and exploration of live-cell single-molecule microscopy images and other hierarchical datasets. **Mol Biosyst** 11(10): 2699-2708. doi: 10.1039/c5mb00321k

76. Stylianidou S, Brennan C, Nissen SB, Kuwada NJ, Wiggins PA (**2016**). SuperSegger: robust image segmentation, analysis and lineage tracking of bacterial cells. **Mol Microbiol** 102(4): 690-700. doi: 10.1111/mmi.13486

77. Ducret A, Quardokus EM, Brun YV (**2016**). MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. **Nat Microbiol** 1(7): 16077. doi: 10.1038/nmicrobiol.2016.77

78. Ursell T, Lee TK, Shiomi D, Shi H, Tropini C, Monds RD, Colavin A, Billings G, Bhaya-Grossman I, Broxton M, Huang BE, Niki H, Huang KC (**2017**). Rapid, precise quantification of bacterial cellular dimensions across a genomic-scale knockout library. **BMC Biol** 15(1): 17. doi: 10.1186/s12915-017-0348-8

79. Sage D, Kirshner H, Pengo T, Stuurman N, Min J, Manley S, Unser M (**2015**). Quantitative evaluation of software packages for single-molecule localization microscopy. **Nat Methods** 12(8): 717-724. doi: 10.1038/nmeth.3442

80. Zawadzki P, Stracy M, Ginda K, Zawadzka K, Lesterlin C, Kapanidis AN, Sherratt DJ (**2015**). The Localization and Action of Topoisomerase IV in Escherichia coli Chromosome Segregation Is Coordinated by the SMC Complex, MukBEF. **Cell Rep** 13(11): 2587-2596. doi: 10.1016/j.celrep.2015.11.034

81. Garza de Leon F, Sellars L, Stracy M, Busby SJW, Kapanidis AN (**2017**). Tracking Low-Copy Transcription Factors in Living Bacteria: The Case of the lac Repressor. **Biophys J** 112(7): 1316-1327. doi: 10.1016/j.bpj.2017.02.028

82. Reyes-Lamothe R, Sherratt DJ, Leake MC (**2010**). Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. **Science** 328(5977): 498-501. doi: 10.1126/science.1185757

83. Cherepanov PP, Wackernagel W (**1995**). Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. **Gene** 158(1): 9-14. doi: 10.1016/0378-1119(95)00193-a

84. Tuson HH, Aliaj A, Brandes ER, Simmons LA, Biteen JS (**2016**). Addressing the Requirements of High-Sensitivity Single-Molecule Imaging of Low-Copy-Number Proteins in Bacteria. **Chemphyschem** 17(10): 1435-1440. doi: 10.1002/cphc.201600035

85. Day RN, Davidson MW (**2009**). The fluorescent protein palette: tools for cellular imaging. **Chem Soc Rev** 38(10): 2887-2921. doi: 10.1039/b901966a

86. Landgraf D, Okumus B, Chien P, Baker TA, Paulsson J (**2012**). Segregation of molecules at cell division reveals native protein localization. **Nat Methods** 9(5): 480-482. doi: 10.1038/nmeth.1955

87. Cranfill PJ, Sell BR, Baird MA, Allen JR, Lavagnino Z, de Gruiter HM, Kremers GJ, Davidson MW, Ustione A, Piston DW (**2016**). Quantitative assessment of fluorescent proteins. **Nat Methods** 13(7): 557-562. doi: 10.1038/nmeth.3891

88. Hellriegel C, Gratton E (**2009**). Real-time multi-parameter spectroscopy and localization in three-dimensional single-particle tracking. **J R Soc Interface** 6 Suppl 1(Suppl_1): S3-14. doi: 10.1098/rsif.2008.0313.focus

89. English BP, Hauryliuk V, Sanamrad A, Tankov S, Dekker NH, Elf J (**2011**). Single-molecule investigations of the stringent response machinery in living bacterial cells. **Proc Natl Acad Sci U S A** 108(31): E365-373. doi: 10.1073/pnas.1102255108

90. Bakshi S, Siryaporn A, Goulian M, Weisshaar JC (**2012**). Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells. **Mol Microbiol** 85(1): 21-38. doi: 10.1111/j.1365-2958.2012.08081.x

91. Uphoff S, Lord ND, Okumus B, Potvin-Trottier L, Sherratt DJ, Paulsson J (**2016**). Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation. **Science** 351(6277): 1094-1097. doi: 10.1126/science.aac9786

92. Stracy M, Kapanidis AN (**2017**). Single-molecule and super-resolution imaging of transcription in living bacteria. **Methods** 120(103-114. doi: 10.1016/j.ymeth.2017.04.001

93. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohana RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J, Darzins A, Klaubert DH, Bulleit RF, Wood KV (**2008**). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. **ACS Chem Biol** 3(6): 373-382. doi: 10.1021/cb800025k

94. Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K (**2003**). A general method for the covalent labeling of fusion proteins with small molecules in vivo. **Nat Biotechnol** 21(1): 86-89. doi: 10.1038/nbt765

95. Chen X, Zaro JL, Shen WC (**2013**). Fusion protein linkers: property, design and functionality. **Adv Drug Deliv Rev** 65(10): 1357-1369. doi: 10.1016/j.addr.2012.09.039

96. Stracy M, Jaciuk M, Uphoff S, Kapanidis AN, Nowotny M, Sherratt DJ, Zawadzki P (**2016**). Single-molecule imaging of UvrA and UvrB recruitment to DNA lesions in living Escherichia coli. **Nat Commun** 7(12568. doi: 10.1038/ncomms12568

97. Stracy M, Lesterlin C, Garza de Leon F, Uphoff S, Zawadzki P, Kapanidis AN (**2015**). Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. **Proc Natl Acad Sci U S A** 112(32): E4390-4399. doi: 10.1073/pnas.1507592112

98. Endesfelder U, Finan K, Holden SJ, Cook PR, Kapanidis AN, Heilemann M (**2013**). Multiscale spatial organization of RNA polymerase in Escherichia coli. **Biophys J** 105(1): 172-181. doi: 10.1016/j.bpj.2013.05.048

99. Cabrera JE, Jin DJ (**2003**). The distribution of RNA polymerase in Escherichia coli is dynamic and sensitive to environmental cues. **Mol Microbiol** 50(5): 1493-1505. doi: 10.1046/j.1365-2958.2003.03805.x

100. Wang Y, Yang Q, Wang Z (**2014**). The evolution of nanopore sequencing. **Front Genet** 5: 449. doi: 10.3389/fgene.2014.00449

101. Rollins GC, Shin JY, Bustamante C, Presse S (**2015**). Stochastic approach to the molecular counting problem in superresolution microscopy. **Proc Natl Acad Sci U S A** 112(2): E110-118. doi: 10.1073/pnas.1408071112

102. Crawford R, Torella JP, Aigrain L, Plochowietz A, Gryte K, Uphoff S, Kapanidis AN (**2013**). Long-lived intracellular single-molecule fluorescence using electroporated molecules. **Biophys J** 105(11): 2439-2450. doi: 10.1016/j.bpj.2013.09.057

103. Grimm JB, English BP, Choi H, Muthusamy AK, Mehl BP, Dong P, Brown TA, Lippincott-Schwartz J, Liu Z, Lionnet T, Lavis LD (**2016**). Bright photoactivatable fluorophores for single-molecule imaging. **Nat Methods** 13(12): 985-988. doi: 10.1038/nmeth.4034

104. Grimm JB, Muthusamy AK, Liang Y, Brown TA, Lemon WC, Patel R, Lu R, Macklin JJ, Keller PJ, Ji N, Lavis LD (**2017**). A general method to fine-tune fluorophores for live-cell and in vivo imaging. **Nat Methods** 14(10): 987-994. doi: 10.1038/nmeth.4403

105. Persson F, Linden M, Unoson C, Elf J (**2013**). Extracting intracellular diffusive states and transition rates from single-molecule tracking data. **Nat Methods** 10(3): 265-269. doi: 10.1038/nmeth.2367

106. Axelrod D, Koppel DE, Schlessinger J, Elson E, Webb WW (**1976**). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. **Biophys J** 16(9): 1055-1069. doi: 10.1016/S0006-3495(76)85755-4

107. Cole NB, Smith CL, Sciaky N, Terasaki M, Edidin M, Lippincott-Schwartz J (**1996**). Diffusional mobility of Golgi proteins in membranes of living cells. **Science** 273(5276): 797-801. doi: 10.1126/science.273.5276.797

108. Garcia V, Phelps SE, Gray S, Neale MJ (**2011**). Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. **Nature** 479(7372): 241-244. doi: 10.1038/nature10515

109. Zhu Z, Chung WH, Shim EY, Lee SE, Ira G (**2008**). Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. **Cell** 134(6): 981-994. doi: 10.1016/j.cell.2008.08.037

110. Mimitou EP, Symington LS (**2008**). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. **Nature** 455(7214): 770-774. doi: 10.1038/nature07312

111. Zou L, Elledge SJ (**2003**). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. **Science** 300(5625): 1542-1548. doi: 10.1126/science.1083430

112. Donnianni RA, Symington LS (**2013**). Break-induced replication occurs by conservative DNA synthesis. **Proc Natl Acad Sci U S A** 110(33): 13475-13480. doi: 10.1073/pnas.1309800110

113. Saini N, Ramakrishnan S, Elango R, Ayyar S, Zhang Y, Deem A, Ira G, Haber JE, Lobachev KS, Malkova A (**2013**). Migrating bubble during break-induced replication drives conservative DNA synthesis. **Nature** 502(7471): 389-392. doi: 10.1038/nature12584

114. Wilson MA, Kwon Y, Xu Y, Chung WH, Chi P, Niu H, Mayle R, Chen X, Malkova A, Sung P, Ira G (**2013**). Pif1 helicase and Poldelta promote recombination-coupled DNA synthesis via bubble migration. **Nature** 502(7471): 393-396. doi: 10.1038/nature12585

115. Liu J, Ede C, Wright WD, Gore SK, Jenkins SS, Freudenthal BD, Todd Washington M, Veaute X, Heyer WD (**2017**). Srs2 promotes synthesis-dependent strand annealing by disrupting DNA polymerase delta-extending D-loops. **Elife** 6. doi: 10.7554/eLife.22195

116. Prakash R, Satory D, Dray E, Papusha A, Scheller J, Kramer W, Krejci L, Klein H, Haber JE, Sung P, Ira G (**2009**). Yeast Mph1 helicase dissociates Rad51-made D-loops: implications for crossover control in mitotic recombination. **Genes Dev** 23(1): 67-79. doi: 10.1101/gad.1737809

117. Ira G, Malkova A, Liberi G, Foiani M, Haber JE (**2003**). Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast. **Cell** 115(4): 401-411. doi: 10.1016/s0092-8674(03)00886-9

118. Wu L, Hickson ID (**2003**). The Bloom's syndrome helicase suppresses crossing over during homologous recombination. **Nature** 426(6968): 870-874. doi: 10.1038/nature02253

119. Ip SC, Rass U, Blanco MG, Flynn HR, Skehel JM, West SC (**2008**). Identification of Holliday junction resolvases from humans and yeast. **Nature** 456(7220): 357-361. doi: 10.1038/nature07470

120. de los Santos T, Hunter N, Lee C, Larkin B, Loidl J, Hollingsworth NM (**2003**). The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. **Genetics** 164(1): 81-94. PMID: 12750322

121. De Muyt A, Jessop L, Kolar E, Sourirajan A, Chen J, Dayani Y, Lichten M (**2012**). BLM helicase ortholog Sgs1 is a central regulator of meiotic recombination intermediate metabolism. **Mol Cell** 46(1): 43-53. doi: 10.1016/j.molcel.2012.02.020

122. Zakharyevich K, Tang S, Ma Y, Hunter N (**2012**). Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. **Cell** 149(2): 334-347. doi: 10.1016/j.cell.2012.03.023

123. Krejci L, Van Komen S, Li Y, Villemain J, Reddy MS, Klein H, Ellenberger T, Sung P (**2003**). DNA helicase Srs2 disrupts the Rad51 presynaptic filament. **Nature** 423(6937): 305-309. doi: 10.1038/nature01577

124. Veaute X, Jeusset J, Soustelle C, Kowalczykowski SC, Le Cam E, Fabre F (**2003**). The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. **Nature** 423(6937): 309-312. doi: 10.1038/nature01585

125. Krejci L, Altmannova V, Spirek M, Zhao X (**2012**). Homologous recombination and its regulation. **Nucleic Acids Res** 40(13): 5795-5818. doi: 10.1093/nar/gks270

126. Symington LS, Rothstein R, Lisby M (**2014**). Mechanisms and regulation of mitotic recombination in Saccharomyces cerevisiae. **Genetics** 198(3): 795-835. doi: 10.1534/genetics.114.166140

127. Brown MS, Bishop DK (**2014**). DNA strand exchange and RecA homologs in meiosis. **Cold Spring Harb Perspect Biol** 7(1): a016659. doi: 10.1101/cshperspect.a016659

128. Heyer WD (**2015**). Regulation of recombination and genomic maintenance. **Cold Spring Harb Perspect Biol** 7(8): a016501. doi: 10.1101/cshperspect.a016501

129. Kowalczykowski SC (**2015**). An Overview of the Molecular Mechanisms of Recombinational DNA Repair. **Cold Spring Harb Perspect Biol** 7(11). doi: 10.1101/cshperspect.a016410

130. San Filippo J, Sung P, Klein H (**2008**). Mechanism of eukaryotic homologous recombination. **Annu Rev Biochem** 77: 229-257. doi: 10.1146/annurev.biochem.77.061306.125255

131. Krogh BO, Symington LS (**2004**). Recombination proteins in yeast. **Annu Rev Genet** 38: 233-271. doi: 10.1146/annurev.genet.38.072902.091500

132. Zelensky A, Kanaar R, Wyman C (**2014**). Mediators of homologous DNA pairing. **Cold Spring Harb Perspect Biol** 6(12): a016451. doi: 10.1101/cshperspect.a016451

133. Liu J, Renault L, Veaute X, Fabre F, Stahlberg H, Heyer WD (**2011**). Rad51 paralogues Rad55-Rad57 balance the antirecombinase Srs2 in Rad51 filament formation. **Nature** 479(7372): 245-248. doi: 10.1038/nature10522

134. Niu H, Klein HL (**2017**). Multifunctional roles of Saccharomyces cerevisiae Srs2 protein in replication, recombination and repair. **FEMS Yeast Res** 17(2). doi: 10.1093/femsyr/fow111

135. Chi P, San Filippo J, Sehorn MG, Petukhova GV, Sung P (**2007**). Bipartite stimulatory action of the Hop2-Mnd1 complex on the Rad51 recombinase. **Genes Dev** 21(14): 1747-1757. doi: 10.1101/gad.1563007

136. Kaniecki K, De Tullio L, Greene EC (**2018**). A change of view: homologous recombination at single-molecule resolution. **Nat Rev Genet** 19(4): 191-207. doi: 10.1038/nrg.2017.92

137. Su GC, Yeh HY, Lin SW, Chung CI, Huang YS, Liu YC, Lyu PC, Chi P (**2016**). Role of the RAD51-SWI5-SFR1 Ensemble in homologous recombination. **Nucleic Acids Res** 44(13): 6242-6251. doi: 10.1093/nar/gkw375

138. Chi P, Van Komen S, Sehorn MG, Sigurdsson S, Sung P (**2006**). Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function. **DNA Repair** 5(3): 381-391. doi: 10.1016/j.dnarep.2005.11.005

139. Chi P, Kwon Y, Seong C, Epshtein A, Lam I, Sung P, Klein HL (**2006**). Yeast recombination factor Rdh54 functionally interacts with the Rad51 recombinase and catalyzes Rad51 removal from DNA. **J Biol Chem** 281(36): 26268-26279. doi: 10.1074/jbc.M602983200

140. Chi P, Kwon Y, Moses DN, Seong C, Sehorn MG, Singh AK, Tsubouchi H, Greene EC, Klein HL, Sung P (**2009**). Functional interactions of meiotic recombination factors Rdh54 and Dmc1. **DNA repair** 8(2): 279-284. doi: 10.1016/j.dnarep.2008.10.012

141. Robertson RB, Moses DN, Kwon Y, Chan P, Chi P, Klein H, Sung P, Greene EC (**2009**). Structural transitions within human Rad51 nucleo-

protein filaments. **Proc Natl Acad Sci U S A** 106(31): 12688-12693. doi: 10.1073/pnas.0811465106

142. Petukhova G, Stratton S, Sung P (**1998**). Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. **Nature** 393(6680): 91-94. doi: 10.1038/30037

143. Petukhova G, Sung P, Klein H (**2000**). Promotion of Rad51-dependent D-loop formation by yeast recombination factor Rdh54/Tid1. **Genes Dev** 14(17): 2206-2215. doi: 10.1101/gad.826100

144. Modesti M, Budzowska M, Baldeyron C, Demmers JA, Ghirlando R, Kanaar R (**2007**). RAD51AP1 is a structure-specific DNA binding protein that stimulates joint molecule formation during RAD51-mediated homologous recombination. **Mol Cell** 28(3): 468-481. doi: 10.1016/j.molcel.2007.08.025

145. Wiese C, Dray E, Groesser T, San Filippo J, Shi I, Collins DW, Tsai MS, Williams GJ, Rydberg B, Sung P, Schild D (**2007**). Promotion of homologous recombination and genomic stability by RAD51AP1 via RAD51 recombinase enhancement. **Mol Cell** 28(3): 482-490. doi: 10.1016/j.molcel.2007.08.027

146. Liang F, Longerich S, Miller AS, Tang C, Buzovetsky O, Xiong Y, Maranon DG, Wiese C, Kupfer GM, Sung P (**2016**). Promotion of RAD51-Mediated Homologous DNA Pairing by the RAD51AP1-UAF1 Complex. **Cell Rep** 15(10): 2118-2126. doi: 10.1016/j.celrep.2016.05.007

147. Zhao W, Steinfeld JB, Liang F, Chen X, Maranon DG, Jian Ma C, Kwon Y, Rao T, Wang W, Sheng C, Song X, Deng Y, Jimenez-Sainz J, Lu L, Jensen RB, Xiong Y, Kupfer GM, Wiese C, Greene EC, Sung P (**2017**). BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. **Nature** 550(7676): 360-365. doi: 10.1038/nature24060

148. Sung P, Robberson DL (**1995**). DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. **Cell** 82(3): 453-461. doi: 10.1016/0092-8674(95)90434-4

149. Smirnova M, Van Komen S, Sung P, Klein HL (**2004**). Effects of tumor-associated mutations on Rad54 functions. **J Biol Chem** 279(23): 24081-24088. doi: 10.1074/jbc.M402719200

150. Raynard S, Sung P (**2009**). Assay for human Rad51-mediated DNA displacement loop formation. **Cold Spring Harb Protoc** 2009(1): pdb prot5120. doi: 10.1101/pdb.prot5120

151. Wright WD, Heyer WD (**2014**). Rad54 functions as a heteroduplex DNA pump modulated by its DNA substrates and Rad51 during D loop formation. **Mol Cell** 53(3): 420-432. doi: 10.1016/j.molcel.2013.12.027

152. Sugiyama T, Zaitseva EM, Kowalczykowski SC (**1997**). A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the Saccharomyces cerevisiae Rad51 protein. **J Biol Chem** 272(12): 7940-7945. doi: 10.1074/jbc.272.12.7940

153. Fien K, Stillman B (**1992**). Identification of replication factor C from Saccharomyces cerevisiae: a component of the leading-strand DNA replication complex. **Mol Cell Biol** 12(1): 155-163. doi: 10.1128/mcb.12.1.155

154. Sneeden JL, Grossi SM, Tappin I, Hurwitz J, Heyer WD (**2013**). Reconstitution of recombination-associated DNA synthesis with human proteins. **Nucleic Acids Res** 41(9): 4913-4925. doi: 10.1093/nar/gkt192

155. Solinger JA, Lutz G, Sugiyama T, Kowalczykowski SC, Heyer WD (**2001**). Rad54 protein stimulates heteroduplex DNA formation in the synaptic phase of DNA strand exchange via specific interactions with the presynaptic Rad51 nucleoprotein filament. **J Mol Biol** 307(5): 1207-1221. doi: 10.1006/jmbi.2001.4555

156. Sung P, Trujillo KM, Van Komen S (**2000**). Recombination factors of Saccharomyces cerevisiae. **Mutat Res** 451(1-2): 257-275. PMID: 10915877

157. Van Komen S, Petukhova G, Sigurdsson S, Stratton S, Sung P (**2000**). Superhelicity-driven homologous DNA pairing by yeast recombination factors Rad51 and Rad54. **Mol Cell** 6(3): 563-572. doi: 10.1016/s1097-2765(00)00055-1

158. Li X, Heyer WD (**2009**). RAD54 controls access to the invading 3'-OH end after RAD51-mediated DNA strand invasion in homologous recombination in Saccharomyces cerevisiae. **Nucleic Acids Res** 37(2): 638-646. doi: 10.1093/nar/gkn980

159. Kwon Y, Daley JM, Sung P (**2017**). Reconstituted System for the Examination of Repair DNA Synthesis in Homologous Recombination. **Methods Enzymol** 591: 307-325. doi: 10.1016/bs.mie.2017.03.021

160. Lydeard JR, Jain S, Yamaguchi M, Haber JE (**2007**). Break-induced replication and telomerase-independent telomere maintenance require Pol32. **Nature** 448(7155): 820-823. doi: 10.1038/nature06047

161. Costantino L, Sotiriou SK, Rantala JK, Magin S, Mladenov E, Helleday T, Haber JE, Iliakis G, Kallioniemi OP, Halazonetis TD (**2014**). Break-induced replication repair of damaged forks induces genomic duplications in human cells. **Science** 343(6166): 88-91. doi: 10.1126/science.1243211

162. Minocherhomji S, Ying S, Bjerregaard VA, Bursomanno S, Aleliunaite A, Wu W, Mankouri HW, Shen H, Liu Y, Hickson ID (**2015**). Replication stress activates DNA repair synthesis in mitosis. **Nature** 528(7581): 286-290. doi: 10.1038/nature16139

163. McVey M, Khodaverdian VY, Meyer D, Cerqueira PG, Heyer WD (**2016**). Eukaryotic DNA Polymerases in Homologous Recombination. **Annu Rev Genet** 50: 393-421. doi: 10.1146/annurev-genet-120215-035243

164. Li X, Stith CM, Burgers PM, Heyer WD (**2009**). PCNA is required for initiation of recombination-associated DNA synthesis by DNA polymerase delta. **Mol Cell** 36(4): 704-713. doi: 10.1016/j.molcel.2009.09.036

165. Li J, Holzschu DL, Sugiyama T (**2013**). PCNA is efficiently loaded on the DNA recombination intermediate to modulate polymerase delta, eta, and zeta activities. **Proc Natl Acad Sci U S A** 110(19): 7672-7677. doi: 10.1073/pnas.1222241110

166. Sebesta M, Burkovics P, Juhasz S, Zhang S, Szabo JE, Lee MY, Haracska L, Krejci L (**2013**). Role of PCNA and TLS polymerases in D-loop extension during homologous recombination in humans. **DNA Repair** 12(9): 691-698. doi: 10.1016/j.dnarep.2013.05.001

167. Sebesta M, Burkovics P, Haracska L, Krejci L (**2011**). Reconstitution of DNA repair synthesis in vitro and the role of polymerase and helicase activities. **DNA Repair** 10(6): 567-576. doi: 10.1016/j.dnarep.2011.03.003

168. Buzovetsky O, Kwon Y, Pham NT, Kim C, Ira G, Sung P, Xiong Y (**2017**). Role of the Pif1-PCNA Complex in Pol delta-Dependent Strand Displacement DNA Synthesis and Break-Induced Replication. **Cell Rep** 21(7): 1707-1714. doi: 10.1016/j.celrep.2017.10.079

169. Burkovics P, Sebesta M, Sisakova A, Plault N, Szukacsov V, Robert T, Pinter L, Marini V, Kolesar P, Haracska L, Gangloff S, Krejci L (**2013**). Srs2 mediates PCNA-SUMO-dependent inhibition of DNA repair synthesis. **EMBO J** 32(5): 742-755. doi: 10.1038/emboj.2013.9

170. Silva S, Altmannova V, Luke-Glaser S, Henriksen P, Gallina I, Yang X, Choudhary C, Luke B, Krejci L, Lisby M (**2016**). Mte1 interacts with Mph1 and promotes crossover recombination and telomere maintenance. **Genes Dev** 30(6): 700-717. doi: 10.1101/gad.276204.115

171. Cejka P, Kowalczykowski SC (**2010**). The full-length Saccharomyces cerevisiae Sgs1 protein is a vigorous DNA helicase that preferentially unwinds holliday junctions. **J Biol Chem** 285(11): 8290-8301. doi: 10.1074/jbc.M109.083196

172. Heyer WD, Ehmsen KT, Solinger JA (**2003**). Holliday junctions in the eukaryotic nucleus: resolution in sight? **Trends Biochem Sci** 28(10): 548-557. doi: 10.1016/j.tibs.2003.08.011

173. Ciccia A, McDonald N, West SC (**2008**). Structural and functional relationships of the XPF/MUS81 family of proteins. **Annu Rev Biochem** 77: 259-287. doi: 10.1146/annurev.biochem.77.070306.102408

174. Mimitou EP, Symington LS (**2009**). Nucleases and helicases take center stage in homologous recombination. **Trends Biochem Sci** 34(5): 264-272. doi: 10.1016/j.tibs.2009.01.010

175. Schwartz EK, Heyer WD (**2011**). Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes. **Chromosoma** 120(2): 109-127. doi: 10.1007/s00412-010-0304-7

176. Wyatt HD, West SC (**2014**). Holliday junction resolvases. **Cold Spring Harb Perspect Biol** 6(9): a023192. doi: 10.1101/cshperspect.a023192

177. Robinson PK (**2015**). Enzymes: principles and biotechnological applications. **Essays Biochem** 59: 1-41. doi: 10.1042/bse0590001

178. McNeil EM, Melton DW (**2012**). DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. **Nucleic Acids Res** 40(20): 9990-10004. doi: 10.1093/nar/gks818

179. Lai X, Broderick R, Bergoglio V, Zimmer J, Badie S, Niedzwiedz W, Hoffmann JS, Tarsounas M (**2017**). MUS81 nuclease activity is essential for replication stress tolerance and chromosome segregation in BRCA2-deficient cells. **Nat Commun** 8: 15983. doi: 10.1038/ncomms15983

180. Wright WD, Ehmsen KT, Heyer WD (**2011**). Assays for structure-selective DNA endonucleases. **Methods Mol Biol** 745: 345-362. doi: 10.1007/978-1-61779-129-1_20

181. Mukherjee S, Wright WD, Ehmsen KT, Heyer WD (**2014**). The Mus81-Mms4 structure-selective endonuclease requires nicked DNA junctions to undergo conformational changes and bend its DNA substrates for cleavage. **Nucleic Acids Res** 42(10): 6511-6522. doi: 10.1093/nar/gku265

182. Ehmsen KT, Heyer WD (**2008**). Saccharomyces cerevisiae Mus81-Mms4 is a catalytic, DNA structure-selective endonuclease. **Nucleic Acids Res** 36(7): 2182-2195. doi: 10.1093/nar/gkm1152

183. Mukherjee S, Wright WD, Ehmsen KT, Heyer WD (**2014**). The Mus81-Mms4 structure-selective endonuclease requires nicked DNA junctions to undergo conformational changes and bend its DNA substrates for cleavage. **Nucleic Acids Res** 42(10): 6511-6522. doi: 10.1093/nar/gku265

184. Wyatt HDM, Laister RC, Martin SR, Arrowsmith CH, West SC (**2017**). The SMX DNA Repair Tri-nuclease. **Molecular Cell** 65(5): 848-860. doi: 10.1016/j.molcel.2017.01.031

185. Zhang XP, Heyer WD (**2011**). Quality control of purified proteins involved in homologous recombination. **Methods Mol Biol** 745: 329-343. doi: 10.1007/978-1-61779-129-1_19

Appendix F Single-Molecule Methods for Nucleotide Excision Repair: Building a System to Watch Repair in Real Time

Review of and protocols for single molecule methods to study DNA repair proteins, originally published in *Methods in Enzymology*.[142] Ref: Kong, M., Beckwitt, E. C., Springall, L., Kad, N. M. & Van Houten, B. Single-Molecule Methods for Nucleotide Excision Repair: Building a System to Watch Repair in Real Time. *Methods Enzymol* **592**, 213-257, doi:10.1016/bs.mie.2017.03.027 (2017).

CHAPTER TEN

Single-Molecule Methods for Nucleotide Excision Repair: Building a System to Watch Repair in Real Time

Muwen Kong[*,†], **Emily C. Beckwitt**[*,†], **Luke Springall**[‡], **Neil M. Kad**[‡], **Bennett Van Houten**[*,†,1]

[*]University of Pittsburgh School of Medicine, Pittsburgh, PA, United States
[†]University of Pittsburgh Cancer Institute, Pittsburgh, PA, United States
[‡]School of Biosciences, University of Kent, Canterbury, United Kingdom
[1]Corresponding author: e-mail address: vanhoutenb@upmc.edu

Contents

Methods in Enzymology, Volume 592
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2017.03.027

Abstract

Single-molecule approaches to solving biophysical problems are powerful tools that allow static and dynamic real-time observations of specific molecular interactions of interest in the absence of ensemble-averaging effects. Here, we provide detailed protocols for building an experimental system that employs atomic force microscopy and a single-molecule DNA tightrope assay based on oblique angle illumination fluorescence microscopy. Together with approaches for engineering site-specific lesions into DNA substrates, these complementary biophysical techniques are well suited for investigating protein–DNA interactions that involve target-specific DNA-binding proteins, such as those engaged in a variety of DNA repair pathways. In this chapter, we demonstrate the utility of the platform by applying these techniques in the studies of proteins participating in nucleotide excision repair.

1. INTRODUCTION

Experiments studying nucleotide excision repair (NER) proteins using optical imaging in our laboratories usually go through three distinct phases: biochemical analysis (Croteau, DellaVecchia, Perera, & Van Houten, 2008; Croteau et al., 2006), atomic force microscopy (AFM) (Wang et al., 2006), and fluorescence single-molecule imaging (Hughes et al., 2013; Kad, Wang, Kennedy, Warshaw, & Van Houten, 2010; Kong et al., 2016). First, proteins should be highly purified and exhibit excellent activity. Purification of these proteins often includes a size-exclusion chromatography step to ensure a homogenous preparation of non-aggregated protein, free of contaminating DNA, which is then examined by a variety of bulk biochemistry methods such as fluorescence anisotropy and electrophoretic mobility shift assays for DNA-binding affinities. These proteins are then imaged alone and complexed with DNA substrates using AFM to assess properties such as homogeneity, stability, stoichiometry (Ghodke et al., 2014; Yeh et al., 2012), specificity, and DNA bend angles (Kong et al., 2016). Finally, the dynamic interactions of these proteins with DNA are visualized with the DNA tightrope assay and fluorescence microscopy (Ghodke et al., 2014; Kad et al., 2010; Kong et al., 2016; Kong & Van Houten, 2016). This chapter first gives detailed protocols on preparing defined DNA substrates for analysis by AFM or our tightrope assay. We then discuss how AFM is used to determine specificity, stoichiometry, and DNA

bend angles. Finally, we end with a description of our optical DNA tightrope flow cell setup with which we can observe quantum dot (Qdot or QD)-labeled proteins using oblique angle illumination on a total internal reflection fluorescence microscope.

2. PREPARATION OF DEFINED LESION SUBSTRATES FOR AFM AND DNA TIGHTROPE ASSAY

To characterize protein–DNA interactions involving proteins that recognize specific targets, DNA sequences or otherwise, it is important to ensure that an optimal number of target sites exist in the DNA substrate against a vast nonspecific background, such that binding events can be observed efficiently. For DNA repair proteins that carry out damage recognition, a common method to globally induce different types of lesions in a random manner is to subject commercially available λ-DNA to physical or chemical manipulations (Kad et al., 2010; Nelson, Dunn, Kathe, Warshaw, & Wallace, 2014). The number of total lesions can be estimated qualitatively for comparison purposes or, in the case of UV-induced photoproducts, explicitly calculated as an average lesion density through quantitative PCR (qPCR) (Furda, Bess, Meyer, & Van Houten, 2012; Meyer et al., 2007). It is also worth noting that UV irradiation of DNA generates 6,4-photoproducts as well as cyclobutane pyrimidine dimers, both of which contribute to the global average lesion density derived from qPCR. Compared to the random distributions of possibly more than one type of lesion generated as briefly described above, a DNA substrate containing site-specific lesion(s) of desired identity offers more control in the sequence context around the lesion site and leads to more predictable binding patterns that may correlate with specific binding events. To this end, we have developed two different strategies for making DNA substrates containing site-specific lesions, suitable for single-molecule AFM and DNA tightrope assays. The first approach, based on the plasmid pSCW01 (Fig. 1A and B) previously used to study DNA mismatch repair, places a 37mer lesion-containing oligonucleotide in a gap created in the plasmid via nicking at four Nt.BstNBI sites (Fig. 1C (i)–(iv)) (Geng et al., 2011; Ghodke et al., 2014). The oligonucleotide containing the defined lesion is sealed into the plasmid by T4 DNA ligase with high efficiency approaching 98%–99% (Fig. 1C (v)). The plasmid can be digested to yield a 538-bp lesion-containing fragment for AFM studies (Fig. 1C; Section 2.7). Alternatively, it is linearized and tandemly ligated (end to end) to form long DNA substrates suitable for the

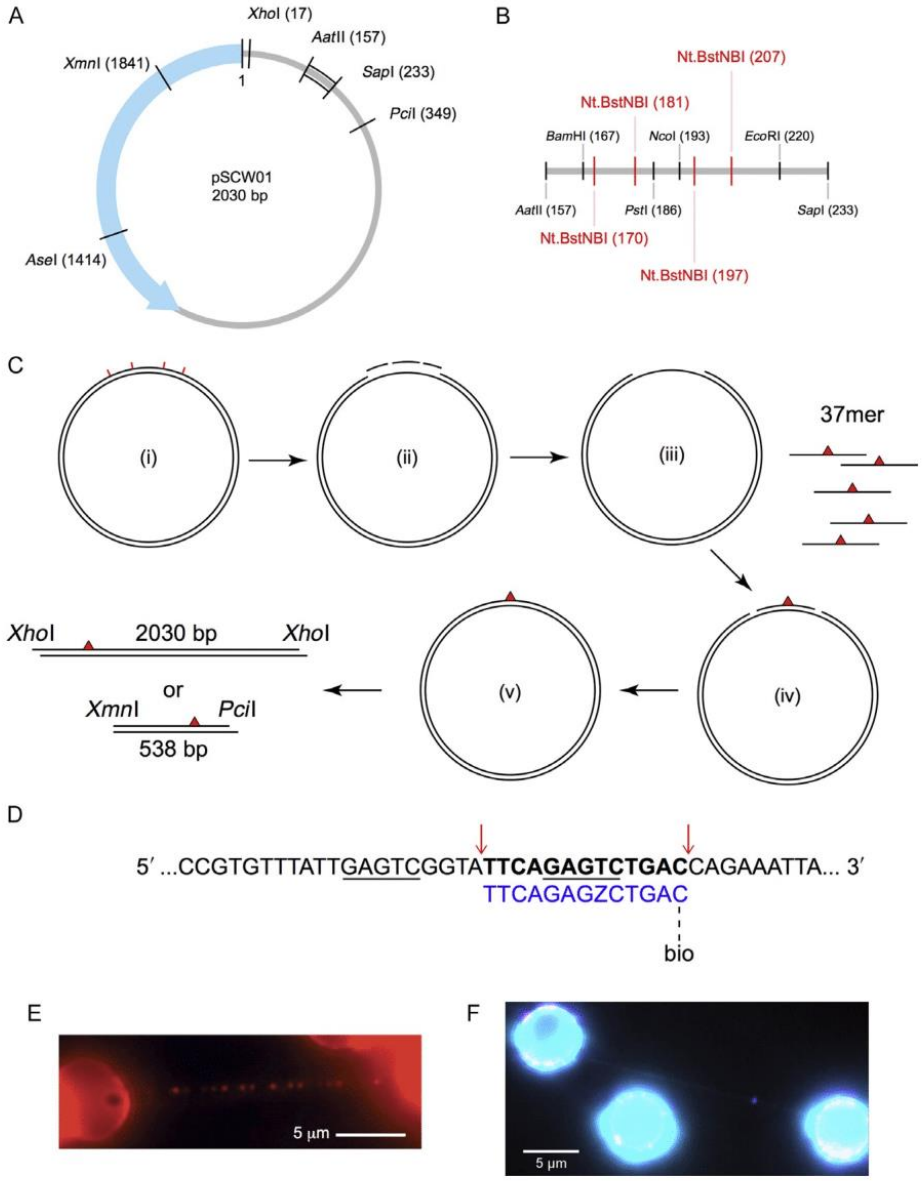

Fig. 1 Design of defined lesion substrate. (A) Map of the pSCW01 plasmid and locations of restriction digest sites. (B) Detailed map of restriction digest and nicking sites for plasmid sequence between the *Aat*II (157) and *Sap*I (233) sites. Nt.BstNBI nicking sites are shown in *red*. (C) Strategy for generating defined lesion substrate based on the pSCW01 plasmid. The plasmid is first nicked by Nt.BstNBI at four different locations (i), which yields three short single-stranded fragments (ii) that are liberated from the plasmid via heating, resulting in a gapped plasmid (iii). 37mer oligonucleotides, each containing a site-specific lesion, are annealed to the gapped plasmids (iv) before the nicks on either side of the oligonucleotides are sealed by overnight ligation (v). The plasmids can now

tightrope assay (Fig. 1C; Sections 2.1–2.6). These defined lesion damage arrays thus contain one site-specific lesion every 2030 bp (Fig. 1E). Another approach inserts an oligonucleotide containing a site-specific lesion into λ-DNA (Fig. 1D; Section 2.8). In this method, λ-DNA is first nicked by Nt.BstNBI at 61 different sites and the shortest single-stranded fragment, between bases 33,778 and 33,791, is then liberated and replaced with a lesion-containing oligonucleotide.

2.1 Growing pSCW01 Plasmid

2.1.1 Equipment

- 37°C shaking incubator
- Laboratory centrifuge

2.1.2 Buffers and Reagents

- *Escherichia coli* transformed with pSCW01 on LB-Amp agar plates
- LB media with 100 μg/mL ampicillin (LB-Amp)

2.1.3 Procedure

1. Pick a single colony from a freshly transformed plate.
2. Inoculate a 2-mL LB-Amp starter culture for 6 h at 37°C.
3. Inoculate 1 L LB-Amp with 1 mL starter culture. Grow for 18 h at 37°C.

---

be linearized by *Xho*I and then tandem-ligated to form long DNA substrates, containing one site-specific damage per 2030 bp, for use in the DNA tightrope assay. Alternatively, the plasmids can be double digested by *Xmn*I and *Pci*I and gel purified to obtain 538 bp fragments, each containing one site-specific damage ~160 bp from the *Pci*I site. (D) Strategy for inserting a damaged oligonucleotide with a biotin conjugate for quantum dot visualization in λ-DNA. The upper λ-DNA sequence is *underlined* at the binding sites for Nt.BstNBI. Cut sites are indicated by *red arrows*, leading to the release of the bolded segment. This is replaced by the lower 5′ phosphorylated oligonucleotide (*blue*) containing damage (*Z* = fluorescein-dT) and biotin-conjugated via TEG at the 3′ end. (E) An array of streptavidin-conjugated quantum dots on a DNA tightrope of a defined lesion substrate containing one site-specific abasic site analog per 2030 bp, each with a proximal biotin marking the site of the lesion. (F) DNA damage (*magenta*) visualized with 655 streptavidin-conjugated quantum dot on a λ-DNA tightrope stained with YOYO-1 (*cyan*). *Panel E: Adapted with permission from Ghodke, H., Wang, H., Hsieh, C. L., Woldemeskel, S., Watkins, S. C., Rapic-Otrin, V., et al. (2014). Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates.* Proceedings of the National Academy of Sciences of the United States of America, 111*(18), E1862–E1871. http://dx.doi.org/10.1073/pnas.1323856111 (fig. 4A).*

4. Centrifuge at 6000 × *g* for 15 min at 4°C to harvest. Store each liter of culture as two pellets.

2.2 Maxiprep of Plasmid DNA

2.2.1 Equipment

- Laboratory centrifuge
- SpeedVac or other vacuum concentrator

2.2.2 Buffers and Reagents

- pSCW01 *E. coli* pellets
- QIAGEN Plasmid Maxi Kit
- Isopropanol
- 70% ethanol

2.2.3 Procedure

1. Resuspend each pellet of culture in 25 mL of buffer P1.
2. Add 25 mL of buffer P2. Incubate at room temperature for 5 min.
3. Add 25 mL of prechilled buffer P3. Mix well.
4. Follow the manufacturer's protocol.
5. Resuspend each DNA pellet in 500 μL of ddH_2O.
6. Concentrate DNA in SpeedVac to ~1 μg/μL.

2.2.4 Notes

1. In step 5, DNA is resuspended in ddH_2O instead of Tris or Tris–EDTA buffer so that samples can be concentrated without affecting concentrations of the buffer components.

2.3 Plasmid DNA Nicking and Oligo Displacement

2.3.1 Equipment

- Heat block or thermocycler

2.3.2 Buffers and Reagents

- Purified plasmid DNA (pSCW01)
- Displacer oligonucleotides (Table 1, IDT)
- Nickase (Nt.BstNBI, 10 U/μL, NEB)
- 10× NEBuffer 3.1 (NEB)

Table 1 Sequences of Oligonucleotides Used in Preparation of Defined Lesion Substrates

Oligonucleotide	Sequence
Displacer1	ATTTGACTCC
Displacer2	CATGGACTCGCTGCAG
Displacer3	GAATGACTCGG
FL37	CCGAGTCATTCC**T**GCAGCGAGTCCATGGGAGTCAAAT
FL37BiodT	CCGAGTCATTCC**T**GCAGCGAGTCCATGGGAGTCAAA/BiodT/
FL13	TTCAGAG**T**CTGAC/BioTEG/

T indicates an internal fluorescein-modified deoxythymidine. /BiodT/ indicates a biotin-modified deoxythymidine. /BioTEG/ indicates a biotin modification attached via a triethylene glycol (TEG) spacer.

2.3.3 Procedure

1. Prepare, in 1 × NEBuffer 3.1 (NEB), purified plasmid DNA (pSCW01) at the final concentration of 400 ng/μL, with 50-fold molar excess of each of the three displacer oligonucleotides (Table 1), and twice the number of units of nickase (Nt.BstNBI, 10 U/μL, NEB) as the amount of plasmid DNA in micrograms. Incubate the reaction at 55°C for 4 h. Before proceeding to the next step, save 1–2 μL of the nicking reaction for diagnostic tests.
2. Inactivate the nicking reaction at 85°C for 10 min before turning off the heat block. Let the heat block cool down to room temperature for approximately 3.5–4 h to allow annealing of displacer oligos with complementary short fragments liberated from plasmids through the nicking reaction. Before proceeding to the next step, save 1–2 μL of the gapped DNA for diagnostic tests.

2.3.4 Notes

1. Start the nicking reaction with at least 50 μg of plasmid DNA for better yield in the next step.
2. During cooling, the excess displacer oligonucleotides capture and anneal to those liberated from the nicking reaction, preventing them from reannealing to the plasmid. These short fragments and oligonucleotides are then removed in the next step.

2.4 PEG Purification of Gapped Plasmid DNA

2.4.1 Equipment

- Heat block or thermocycler
- Benchtop centrifuge
- Nanodrop or other UV–vis spectrophotometer
- Standard agarose gel electrophoresis equipment

2.4.2 Buffers and Reagents

- 2 × PEG solution (26% polyethylene glycol, MW 8000 and 20 m*M* $MgCl_2$)
- 70% ethanol
- 1 × TE buffer (10 m*M* Tris–HCl, pH 8.0, 1 m*M* EDTA)
- Restriction enzymes (*Pst*I and *Nco*I, NEB)

2.4.3 Procedure

1. Pool and transfer now gapped DNA plasmids to *new* Eppendorf tubes, each containing no more than 500 μL in volume. Add equal volume of 2 × PEG solution to each tube and mix well.
2. Centrifuge at 4°C for 1 h at the maximum speed (14,800 rpm) on a benchtop centrifuge.
3. Carefully remove the supernatant from each tube. Precipitated DNA should have formed a thin film stuck on the side of the tube. Using a pipette, wash the side wall with 500 μL of 70% ethanol. The white film of DNA should peel off and settle to the bottom of the tube.
4. Centrifuge and collect the DNA pellet at 4°C for 15 min at the maximum speed (14,800 rpm) on a benchtop centrifuge.
5. Carefully remove the supernatant from each tube without disturbing the DNA pellet at the bottom.
6. Air dry the tube and the pellet before resuspending the pellet in 200 μL of ddH_2O.
7. Dilute 1 μL of the purified gapped plasmid DNA in 20 μL of 1 × TE buffer and measure the DNA concentration at A_{260} using a UV–vis spectrophotometer (NanoDrop 2000, Thermo Scientific).
8. Save 1–2 μL of the purified gapped plasmid DNA for diagnostic tests.
9. Test for completeness of nicking and gapping reactions by setting up restriction digests of samples saved previously after nicking and gapping reactions. Purified pSCW01 plasmids should be used as a positive control. Restriction enzymes (*Pst*I and *Nco*I, NEB) target the sequence that is nicked and/or liberated after nicking, and therefore will not incise the gapped plasmid DNA. Typical reactions contain 100–200 ng of nicked or gapped plasmid DNA and 5 U of restriction enzyme in 20 μL of appropriate reaction buffer and are incubated at 37°C for 2 h. Run all digested reactions and undigested controls on a 1% agarose gel (Fig. 2A).

2.4.4 Notes

1. During the resuspension step, it may be helpful to heat the tube at 55°C for 10 min to help resolubilize the DNA.

2.5 Annealing and Ligation of 37mer Oligo

2.5.1 Equipment

- Heat block
- Thermocycler or heat block in cold room or fridge
- Standard agarose gel electrophoresis equipment

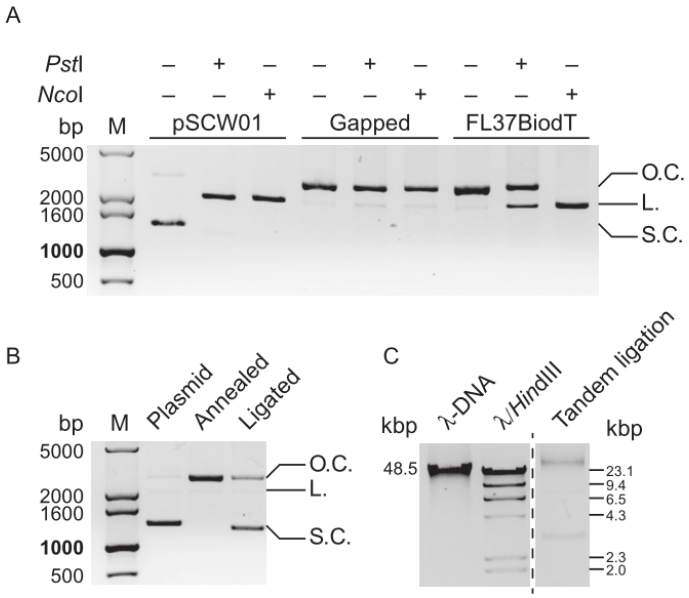

Fig. 2 Diagnostic agarose gels for preparation of defined lesion substrates. (A) 1% agarose gel of diagnostic restriction digests of pSCW01 plasmid, gapped plasmid, and gapped plasmid with FL37BiodT annealed. *Pst*I and *Nco*I, whose restriction sites are within the 37-base gap, do not linearize the gapped plasmid. With FL37BiodT annealed in the gap, *Nco*I linearizes the now nicked plasmid. Restriction digest by *Pst*I on the FL37BiodT-annealed plasmid is hindered due to the presence of the fluorescein in the *Pst*I restriction site. *L.*, linearized plasmid; *O.C.*, open circle, nicked or gapped plasmid; *S.C.*, supercoiled plasmid. (B) 1% agarose gel of pSCW01 plasmid, FL37BiodT-annealed plasmid, and FL37BiodT-ligated plasmid. The reappearance of the supercoiled band in the ligated plasmid lane indicates completion of the ligation reaction. *L.*, linearized plasmid; *O.C.*, open circle, nicked or gapped plasmid; *S.C.*, supercoiled plasmid. (C) 0.8% agarose gel of tandem-ligated the FL37-containing defined lesion substrate with full-length λ-DNA and λ-DNA *Hin*dIII digest fragments as size markers.

- Standard denaturing polyacrylamide electrophoresis equipment
- Typhoon scanner (GE Healthcare)

2.5.2 Buffers and Reagents

- 10× NEBuffer 2.1 (NEB)
- Lesion-containing 37mer oligonucleotides (Table 1, IDT)
- Fluorescently labeled 37mer and 50mer oligonucleotides (IDT)
- 100 m*M* ATP solution
- T4 DNA ligase (NEB)
- Restriction enzymes (*Pst*I, *Nco*I, *Eco*RI, and *Aat*II, NEB)
- 2× denaturing sample loading buffer (NEB)

2.5.3 Procedure

1. Fill the gap by annealing a 37mer oligonucleotide that contains a lesion of choice. Always carry out the annealing and ligation steps for the lesion-containing oligo in parallel to the same experiments using a fluorescein-labeled 37mer (FL37, Table 1), which can be later used to check annealing and ligation reactions.
2. Set up annealing reactions in 1× NEBuffer 2.1 (NEB), containing 400 n*M* gapped plasmids and threefold molar excess of 37mer lesion-containing oligonucleotides (and in parallel, FL37). Incubate at 85°C for 10 min before turning off the heat block. Let the heat block cool down to room temperature for approximately 3.5–4 h to allow annealing of 37mer oligonucleotides. Save 1–2 μL of the annealed plasmid DNA for diagnostic tests.
3. Set up ligation reaction to seal the 5′- and/or 3′-nicks that remain after annealing. To the annealing reaction, add ATP and T4 DNA ligase (2000 U/μL, NEB) to a final concentration of 8 m*M* and 20 U/μL, respectively. Incubate at 16°C for 18 h.
4. Inactivate the ligation reaction at 65°C for 10 min before turning off the heat block. Let the heat block cool down slowly to room temperature.
5. Save 1–2 μL of each ligation reaction for diagnostic tests.
6. To test for completeness of annealing reaction, set up restriction digest reactions of saved sample of annealed plasmid DNA with restriction enzymes that target the sequence in the annealed oligo. Incubate 100–200 ng of annealed plasmids with 5 U of restriction enzymes (*Pst*I or *Nco*I, NEB) in 20 μL reaction volume in appropriate buffers for 2 h at 37°C and run with undigested control on 1% agarose gel (Fig. 2A).
7. To test for completeness of the ligation reaction, set up restriction digest reactions of the saved ligated plasmids with FL37 in the gap. Prepare single digestions of the sample with either *Eco*RI or *Aat*II, as well as a double digestion with both enzymes. Incubate 20 μL reactions containing 100–200 ng plasmids and 5 U of restriction enzyme(s) in appropriate buffer at 37°C for 2 h. To each 5 μL of digested samples and undigested control, as well as 2 μL of 25 n*M* fluorescein-labeled oligonucleotides of appropriate lengths (37mer and 50mer), add equal volume of 2× denaturing sample loading buffer. Heat all samples at 90°C for 5 min and chill on ice immediately. Load these samples on a prerun 10% denaturing polyacrylamide gel. Ensure that the gel runs hot to the touch to prevent reannealing of single-stranded DNA and image on a fluorescence scanner (Typhoon 9400, GE Healthcare). Lengths of the

diagnostic restriction digests will vary depending on whether the 5′- or the 3′-nick was sealed. We normally observe >98% ligation of both ends of the modified 37mer.

2.5.4 Notes

1. Ideally, steps in the protocol from nicking plasmids to annealing of 37mer oligonucleotides should be completed in 1 day, with the 18-h ligation setup to take place overnight. This is so that the time that plasmids remain gapped, during which they are presumably the most fragile, is minimized. However, if necessary, purified gapped plasmids can be stored overnight at 4°C without significant adverse effects on the quality of the entire preparation.
2. Ligation reaction can also be confirmed by comparing overnight-ligated plasmids to those before ligation. A supercoiled band similar to that seen in purified plasmids should reappear after ligation (Fig. 2B).

2.6 Linearization and Tandem Ligation

2.6.1 Equipment

- Heat block or thermocycler
- Standard agarose gel electrophoresis equipment

2.6.2 Buffers and Reagents

- 10× NEBuffer 2.1 (NEB)
- 50 m*M* EDTA
- Restriction enzyme (*Xho*I, NEB)
- T4 DNA ligase (NEB)
- 2× Quick Ligation Reaction Buffer (NEB)
- Dry ice
- λ-DNA and λ-DNA *Hin*dIII digest fragments (NEB)

2.6.3 Procedure

1. Linearize ligated plasmids by incubating them with twice the number of units of *Xho*I (20 U/μL, NEB) as the amount of DNA in micrograms at 37°C for 2 h. Adjust the final concentration of NEBuffer 2.1 (NEB) to 1× with 10× stock if necessary.
2. Heat inactivate *Xho*I by incubating at 80°C for 20 min. Turn off heat block and allow it to cool down slowly to room temperature. Linearized plasmids can be stored at –20°C.

3. For tandem (end-to-end) ligation of linearized plasmids: incubate 1 μg of plasmids with 2 μL of T4 DNA ligase (2000 U/μL, NEB) in 1 × Quick Ligation Reaction Buffer (NEB) in a total reaction volume of 20 μL at room temperature for 15 min.
4. At the end of the ligation, save 2 μL of the reaction and stop the reaction by adding 1 μL of 50 m*M* EDTA. Stop the rest of the reaction (18 μL) by placing the ligation reaction tube on dry ice till frozen. Ligation products can be kept for short-term storage at –20°C.
5. To check the efficiency of tandem ligation, run the saved sample from the step above on 0.8% agarose gel with full-length λ-DNA (NEB) and λ-DNA *Hin*dIII digest fragments (NEB) as standards (Fig. 2C). Tandem-ligation products should be at least the same length as the longest λ-DNA *Hin*dIII digest fragment (~23,000 bp), preferably equal to or longer than λ-DNA (~48,000 bp).

2.7 Preparation of DNA Substrate for AFM

2.7.1 Equipment

- Heat block or thermocycler
- Standard agarose gel electrophoresis equipment
- UV transilluminator
- Benchtop centrifuge
- Nanodrop or other UV–vis spectrophotometer
- SpeedVac or other vacuum concentrator

2.7.2 Buffers and Reagents

- Restriction enzymes (*Xmn*I and *Pci*I, NEB)
- 10 × NEBuffer 2.1 (NEB)
- Agarose gel purification kit (Wizard SV Gel and PCR Clean-Up System, Promega)
- PCR purification kit (QIAquick PCR Purification Kit, Qiagen)
- AFM water: autoclaved nuclease-free ddH_2O, 0.02 μm filtered

2.7.3 Procedure

1. Set up a double digest with restriction enzymes (*Xmn*I and *Pci*I, NEB) in appropriate buffer (1 × NEBuffer 2.1, NEB). Use twice the number of units of each restriction enzyme as the amount of annealed and ligated lesion-containing plasmid DNA in micrograms. Incubate the reaction at 37°C for 4 h.

2. Inactivate the digestion reaction at 80°C for 20 min before turning off the heat block. Let the heat block cool down slowly to room temperature. Run a small sample of the digested product on 1% agarose gel to ensure that digestion was complete.
3. Run the rest of digestion reaction on 1% agarose gel. Excise the band of appropriate size from gel and extract DNA with a commercial gel purification kit per manufacturer's protocol. See Section 2.7.4 for notes on avoiding UV damage.
4. Purify gel-extracted DNA one more time with a commercial PCR purification kit (QIAquick PCR Purification Kit, Qiagen) per manufacturer's protocol to ensure complete removal of restriction enzymes from the desired DNA fragments. The final elution of DNA should be carried out in AFM water. Measure DNA concentration at A_{260} using a UV–vis spectrophotometer (NanoDrop 2000, Thermo Scientific).
5. In a vacuum concentrator (SpeedVac DNA120, Thermo Scientific), concentrate DNA sample to desired concentration appropriate for AFM-binding experiments (~200–300 n*M*). DNA can be kept at 4°C for immediate use, or –80°C for long-term storage.

2.7.4 Notes

1. When excising gel bands on the UV transilluminator, it is important to minimize the bands' exposure to UV as UV light could induce additional undesired photoproducts in DNA. To do so, load in a separate lane a small amount of digested DNA for visualization purpose only and shield the bulk of the DNA sample in gel from UV with aluminum foil.
2. Two-step purification (gel extraction and PCR purification kits) should remove all DNA-bound restriction enzymes from the sample. However, if proteins are found bound to DNA upon quality check under AFM, additional rounds of PCR purification may be needed at the cost of slight loss of DNA sample. Additionally, it may be necessary to do a phenol–chloroform extraction and ethanol precipitation to get rid of stubborn proteins.
3. It may be desirable to aliquot purified DNA sample into single-use tubes and store at –80°C to avoid repeated freeze–thaw cycles.

2.8 Defined Lesion Substrates Based on λ-DNA

2.8.1 Equipment

- Heat block

2.8.2 Buffers and Reagents

- Nickase (Nt.BstNBI, 10 U/μL, NEB)
- 10 × NEBuffer 3.1 (NEB)
- λ-DNA (NEB)
- T4 DNA ligase (1 U/μL, NEB)
- 100 m*M* ATP solution
- 1 μ*M* 13mer oligonucleotide with site-specific damage at position 8 and a 3′ biotin modification via a TEG linker (Table 1)
- Qdot Streptavidin Conjugate (Thermo Scientific)
- YOYO-1 dye (Thermo Scientific)

2.8.3 Procedure

1. Prepare the nicking reaction using NEBuffer 3.1, 5 μg of λ-DNA and 2 U of enzyme; incubate at 55°C for 2 h.
2. Digestion of λ-DNA with the single-stranded nickase will create numerous nicks with which only one pair will be close enough together to generate an oligonucleotide fragment with a near room temperature melting point, regions 33,778–33,791 of λ-DNA (Fig. 1D).
3. Incubate with a 10-fold excess of damage-containing oligonucleotide (FL13, Table 1) at 55°C for 10 min.
4. Allow the solution to cool to room temperature.
5. Perform the ligation with 1 U of T4 DNA ligase and 1 m*M* ATP at room temperature overnight.
6. Removal of DNA ligase can be achieved using phenol:chloroform extraction (Sambrook, Fritsch, & Maniatis, 1989).
7. The lesion-containing DNA is ready to be used for DNA tightropes.
8. The DNA can be stored at 4°C for use within a day or two, for longer storage −20°C is preferred.
9. To visualize the damage site located 5 bases from the biotin, add 10 n*M* streptavidin-conjugated Qdots into a flow cell and incubate for 15 min. This can be combined with 100 n*M* YOYO-1 dye to visualize the DNA simultaneously (Fig. 1F).

2.8.4 Notes

1. For longer tightropes, DNA can be concatemerized (Springall, Inchingolo, & Kad, 2016).
2. This procedure is based on the method of Tafvizi, Huang, Fersht, Mirny, and van Oijen (2011).

3. ATOMIC FORCE MICROSCOPY

AFM provides a topographical view of protein–DNA interactions (Fig. 3). Three major sets of data can be obtained from a single protein–DNA experiment: protein specificity for site-specific lesions, as determined by its binding position on a DNA substrate (Fig. 4B and F); the bend angle of DNA at points of specific and nonspecific protein binding or otherwise (Fig. 4C and G); and the stoichiometry of protein binding to DNA substrates as determined by the volume of the complex (Fig. 4A, D, and E). The steps needed to acquire these data are outlined below. The overall process involves setting up a protein–DNA-binding reaction and depositing the sample onto atomically smooth mica (Section 3.1), imaging with an atomic force microscope (Section 3.2), and analyzing data (Section 3.3).

Binding reactions are set up using purified proteins and DNA substrates 500–600 bp in length. The process described in Section 2.7 produces a 538-bp DNA duplex with a single site-specific lesion, positioned at 30% the contour length. Empirically, substrates of this size are ideal for AFM because they are long enough to allow for precise assessment of protein-binding positions and DNA bend angles, but short enough such that a large number of molecules can be captured in a single $1 \times 1\,\mu m$ field without excessive overlap and convolution.

Many protocols for AFM take advantage of the chemical properties of mica. First, mica exists in sheets that can be easily cleaved. Freshly cleaved mica is an atomically smooth surface, ideal for AFM imaging, as it will not contribute to the landscape being imaged. Second, the surface of freshly cleaved mica has a negative charge, which may be useful for studying certain positively charged particles. However, when studying protein–DNA interactions, mica can be treated with divalent and/or monovalent cations; we use a combination of sodium and magnesium salts in our deposition buffer. This confers a positive charge to the mica surface that will attract the negative phosphate backbone of DNA and enhance sample adhesion (Hansma & Laney, 1996; Vesenka et al., 1992).

Finally, the atomic force microscope scans the samples on mica to produce topographical data. Suspension of the microscope with bungee cords provides some protection from interfering vibrations (Fig. 3A). In AFM tapping mode (Fig. 3B), a cantilever (with probe tip at the end) is driven to oscillate vertically near its resonance frequency. The AFM scanner allows the probe to track a sample field in the X–Y dimensions. In tapping mode,

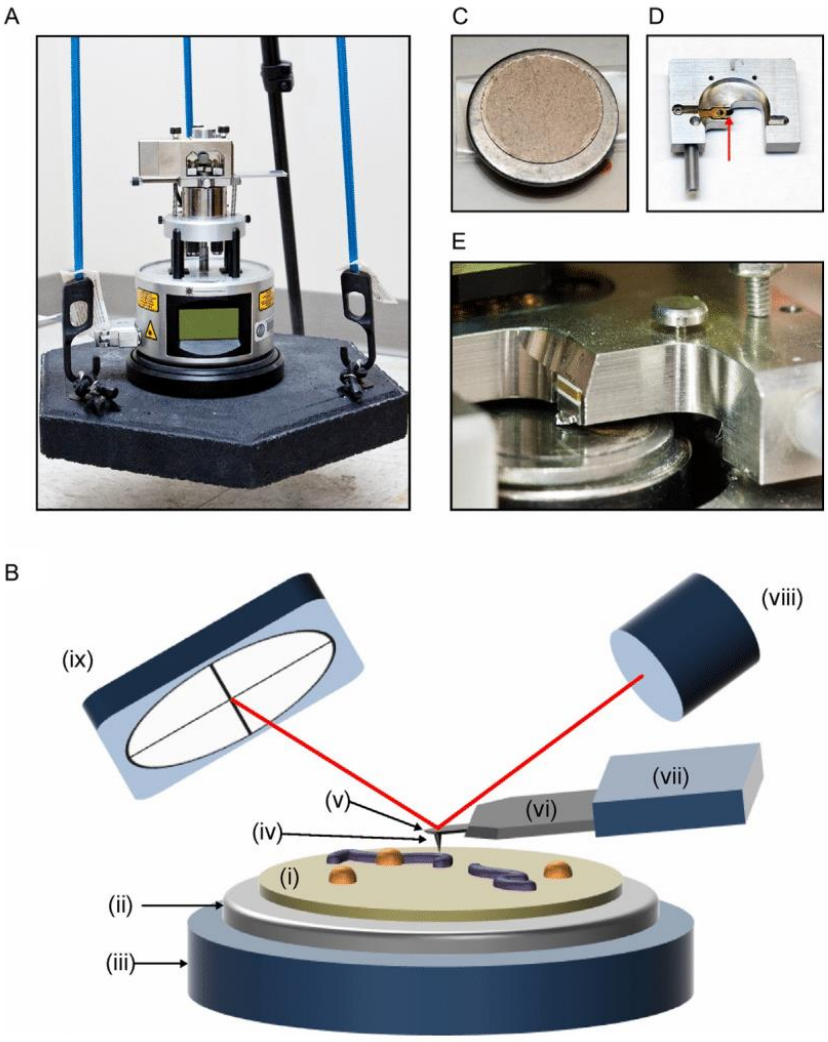

Fig. 3 Atomic force microscopy setup. (A) diMultiMode V atomic force microscope by Veeco. For noise isolation, the AFM is placed on a heavy platform suspended by elastic bungee cords that are secured to a tripod. (B) Schematic of AFM tapping mode in air (not to scale). A protein–DNA sample on mica (i) glued to a metal disc (ii) is placed on the AFM scanner (iii). The probe tip (iv) scans across the sample to generate AFM data. The tip is located at the end of a cantilever (v), which is attached to a support chip (vi) and held by the probe holder (vii). A laser (viii) is reflected off the cantilever and onto a photodetector (ix). Deflection of the cantilever induced by the sample surface changes the path of the laser beam and provides topographical information about the sample (not shown). (C) 12-mm mica chip glued to metal disc. (D) Probe holder with probe (*red arrow*) for tapping in air. (E) Close-up of probe holder and probe installed above mica on scanner.

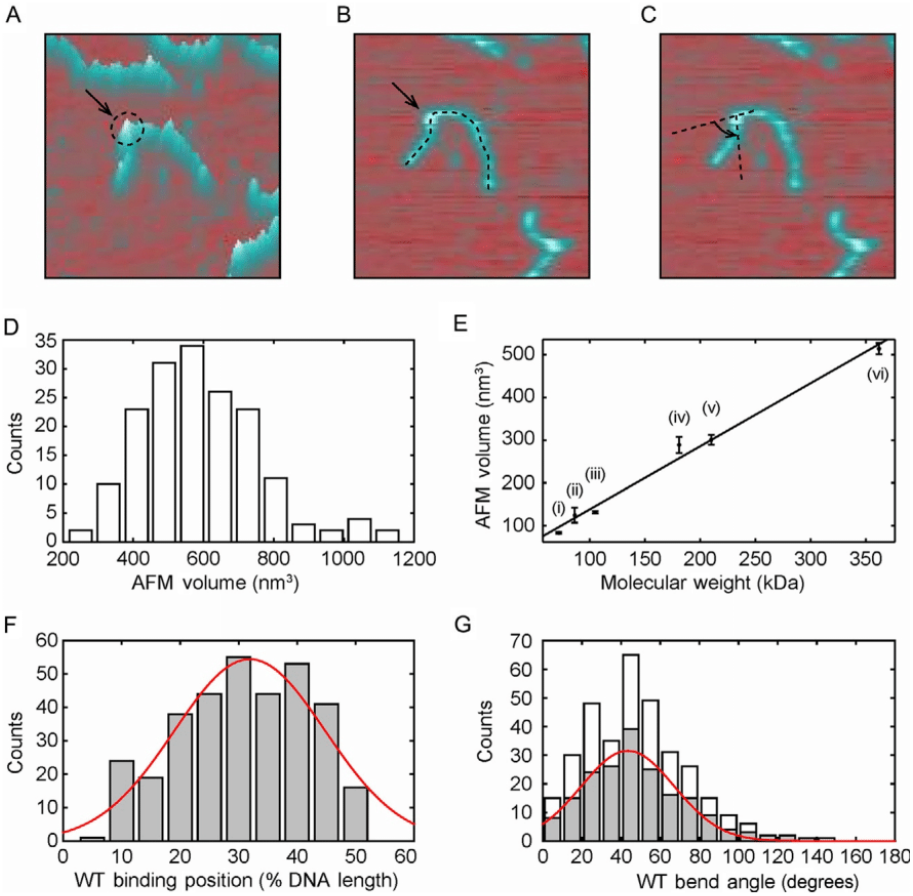

Fig. 4 AFM imaging of protein volume, position, and bend angle. (A)–(C) AFM images of each protein-binding event on DNA can be used to extract information on the protein volume, binding position, and DNA bend angle, respectively. (D) Histogram of UV-DDB K244E mutant volumes on DNA ($n = 171$). (E) Calibration curve relating the molecular weight of a complex to its measured AFM volume, mean ± SD of three separate determinations. The curve was generated using the following proteins in solution: (i) Pot1 monomer (65 kDa), (ii) PcrA monomer (86.4 kDa), (iii) UvrA monomer (105 kDa), (iv) Taq MutS dimer (181 kDa), (v) UvrA dimer (210 kDa), and (vi) Taq MutS tetramer (362 kDa). Linear fit to the data yields $V(\text{nm}^3) = 1.471\ \text{MW (kDa)} - 7.294$ with $R^2 = 0.9886$. (F) Histogram and Gaussian fitting (*red curve*) of wild-type Rad4–Rad23-binding positions (32% ± 13%, $n = 335$) on 538 bp DNA fragment in terms of percentage of total contour length measured from one end. (G) Histogram of DNA bend angles at all internal wild-type Rad4–Rad23-binding sites (*white*, $n = 335$). The histogram (*gray*) and Gaussian fitting (*red curve*) show DNA bend angles (43 ± 24°, $n = 189$) at specific binding events (proteins bound between 20% and 40%). *Panel (D): Adapted with permission from Ghodke, H., Wang, H., Hsieh, C. L., Woldemeskel, S., Watkins, S. C., Rapic-Otrin, V., et al. (2014). Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates.* Proceedings of the National

the oscillation amplitude is kept constant, but interaction with the sample surface causes deflection of the cantilever and alters the path of the reflected laser beam. Three-dimensional images are captured and can be analyzed using various computer programs.

3.1 Binding Reaction and Sample Preparation

3.1.1 Equipment

- Heat block
- Mica (SPI) fixed to metal disks with low-melt glue (SPI)
- Forceps
- Compressed nitrogen gas
- Scotch tape (3M)

3.1.2 Buffers and Reagents

- Purified protein, ~5 μ*M*
- DNA substrate, ~200 n*M* (Section 2.7)
- Protein-binding buffer (Note 2)
- Deposition buffer: 25 m*M* HEPES, pH 7.5, 25 m*M* NaOAc, 10 m*M* $Mg(OAc)_2$, 0.02 μm filtered
- AFM water: autoclaved nuclease-free ddH_2O, 0.02 μm filtered

3.1.3 Procedure

1. Set up binding reaction. This will vary depending on the protein and DNA being studied. In general, a 10-μL reaction can be prepared with 100 n*M* DNA substrate and 500 n*M* protein in binding buffer. Allow the reaction to proceed for 30 min at room temperature.

Academy of Sciences of the United States of America, 111*(18), E1862–E1871.* http://dx.doi.org/10.1073/pnas.1323856111 *(fig. 5E). Panel (E): Adapted with permission from Ghodke, H., Wang, H., Hsieh, C. L., Woldemeskel, S., Watkins, S. C., Rapic-Otrin, V., et al. (2014). Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates.* Proceedings of the National Academy of Sciences of the United States of America, 111*(18), E1862–E1871.* http://dx.doi.org/10.1073/pnas.1323856111*, fig. S6D. Panel (F): Adapted with permission from Kong, M., Liu, L., Chen, X., Driscoll, K. I., Mao, P., Bohm, S., et al. (2016). Single-molecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process.* Molecular Cell, 64*(2), 376–387.* http://dx.doi.org/10.1016/j.molcel.2016.09.005 *(fig. 5A). Panel (G): Adapted with permission from Kong, M., Liu, L., Chen, X., Driscoll, K. I., Mao, P., Bohm, S., et al. (2016). Single-molecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process.* Molecular Cell, 64*(2), 376–387.* http://dx.doi.org/10.1016/j.molcel.2016.09.005 *(fig. 5B).*

2. During reaction, heat the required volume (~200 μL) of deposition buffer at 65°C for 15–20 min. After heating, vortex buffer and spin down briefly. Allow to cool back to room temperature before setting up dilutions in step 4.
3. During reaction, and while deposition buffer is preheating, cleave mica using scotch tape. A razor blade may be used to make a shallow cut across the edge of the mica as a starting point for peeling. Smooth tape over the surface of the mica and, gripping the metal disk with forceps, pull the tape back. Check surface for uneven cleavage and repeat if necessary. A mica chip glued on a metal disc is shown in Fig. 3C.
4. When steps 1–3 are complete, set up depositions one at a time, in order to minimize time sample is spent in deposition buffer (Note 3). Add 1 μL of the reaction to 24 μL of deposition buffer and mix gently. Transfer all 25 μL of the diluted reaction onto the mica (Note 4); be careful not to touch the surface with the pipette tip. Immediately after depositing the droplet, gently rock the mica back and forth and swirl to distribute the sample evenly on the surface. Do this for 30 s and immediately begin step 5.
5. Aspirate 1000 μL of AFM water in a micropipette, dispense approximately 200 μL onto the mica surface, and flick water into sink. Repeat until you have used all the water (~5 washes total).
6. Dry the mica under a gentle stream of N_2 gas. Push the liquid off the mica and onto a paper towel. Be careful of the air stream such that water droplets run down and off the surface, but are not allowed to come back up.

3.1.4 Notes

1. It is important that the purified protein is very clean and is stored in a buffer that does not include BSA, as this will interfere with imaging and analysis in the following sections.
2. Protein-binding buffer will vary depending on the specific reaction being studied. 50 m*M* HEPES (pH 7.5) and 150 m*M* NaCl are a good starting point.
3. These steps would be easiest with three hands, but they are manageable with a little forethought. We recommend setting up the "wash station" prior to beginning the depositions: aspirate 1000 μL of AFM water and leave pipette by N_2 tank, along with some paper towels laid on the bench to collect runoff. Then, make the dilution and aspirate the sample. Carefully set the pipette on the bench while picking up the mica with forceps.

Then, the operator can transfer the forceps/mica to their nondominant hand and dispense the sample using their dominant hand.

4. Sample concentration and volume both affect distribution on the mica. Typically, depositing 25 μL of a 1:25 dilution (for a final concentration of 20 n*M* protein and 4 n*M* DNA) of the reaction results in favorable sample distribution without overcrowding. However, some optimization may be required. We suggest setting up multiple depositions to test these factors.

3.2 Imaging With AFM

3.2.1 Equipment

- Atomic force microscope (Veeco, diMultiMode V, or other/newer models)
- AFM controller (Bruker, NanoScope V)
- NanoScope 9.0 software (earlier versions will also work)
- Tripod with bungees or air table to protect AFM from environmental vibrations
- Probes for tapping in air (Nanosensors PointProbe® Plus)

3.2.2 Procedure

1. Turn on the AFM controller.
2. Open the NanoScope software and begin a Tapping in Air protocol. Select a Capture Directory for files to be saved. Basic steps of the procedure are outlined on the left of the window.
3. Begin with *Setup*. Enter probe information if desired.
4. On the microscope itself, switch the mode to *AFM/LFM*. Insert a fresh probe into the probe holder (Fig. 3D). Adjust the laser and mirror positions for the maximum signal intensity.
5. Remove the AFM head by releasing the springs on either side and place the mica onto the magnetic sample pedestal. Carefully, replace the AFM head and reattach the springs. The probe is now positioned over the sample on the scanner (Fig. 3E).
6. Lower the tip to ~50–100 μm above the mica surface. When using the NanoSensors PointProbe® Plus, this can be estimated as roughly half the length of the cantilever that is visible.
7. Use the AFM knobs to adjust the laser position on the detector. The display on the AFM should read as close to 0 as possible for both vertical and horizontal differences.

8. Switch the AFM to *Tapping/TFM* mode and carefully transfer the AFM to the bungee setup (Fig. 3A).
9. Still on the *Setup* step in NanoScope, press *Auto Tune*. Verify that the Drive Amplitude is less than 100 mV. It may be necessary to use *Manual Tune* to achieve appropriate settings.
10. Click on the next step in NanoScope: *Check Parameters*. Begin with the following settings: scan size: 0.00 nm; aspect ratio: 1.00; *X* offset: 0.00 nm; *Y* offset: 0.00 nm; scan rate: 3.26 Hz; Samples/Line: 512; Lines: 512.
11. Click on the next step in NanoScope: *Engage*. The tip will lower toward the mica until it engages and begins tracking the sample. Ensure you are scanning in the height channel. Because the scan size is set to 0, the surface should appear completely flat. Verify that the Trace and Retrace curves are both sufficiently flat.
12. Click on the *Withdraw* step in NanoScope and then return to *Setup*. Repeat steps 9–11.
13. Increase the scan size to 1000 nm. Press the *Frame Up* or *Frame Down* arrows to begin at the bottom or top of the field, respectively.
14. While scanning, capture the current field by pressing the camera button (Capture). The status bar will read "Capture: On" for the duration of the scan, and "Capture: Done" when it is complete and the file has been saved to the Capture Directory. The status bar will read "Capture: Next" if parameters have been changed within the current scan, such as scan size or offset.
15. After each image, change the *X* and *Y* offsets to move to a new area on the mica and capture a new image. We suggest moving by 1.1 μm each time to account for drift and avoid redundancy.
16. Open raw 001 files in NanoScope Analysis. Flatten the images using the *Flatten* tool. Select first- (line by line) or second (to correct for bowing effect)-order flattening and press *Execute*. Adjust data scales and colors as desired. Save changes.
17. To export BMP files of the height images, select the desired files in the Browse Menu. *Right-click* > *Export…* > *bmp*.
18. When imaging is complete, withdraw the trip (press *Withdraw* several times) and remove the AFM from the bungee setup. Close the NanoScope software and then shut off the controller.

3.2.3 Notes

1. When inserting probes and samples, be careful to keep the tip position high off the surface to prevent accidental damage.

2. Analysis can be done on any of the following file formats: BMP, TIFF, and JPEG.

3.3 Data Analysis

3.3.1 Equipment

- ImageJ (https://imagej.nih.gov/ij/, NIH)
- Image SXM
- Microsoft Excel
- GraphPad Prism 7 or other data analysis program

3.3.2 Procedure

1. Open the image file (Section 3.2.2, step 17) with ImageJ.
2. Label protein–DNA complexes using the Text Tool. Use the following criteria to ensure that all data points will provide reliable information:
   a. The protein–DNA complex is larger than unbound DNA (typically in both height and area).
   b. Entire DNA molecule is visible in the image. It must not continue past the edge of the image nor overlap with other molecules.
   c. DNA is the correct length. This can be judged initially by eye, and again when measuring the contour length (step 3). DNA molecules within 10% of the expected length can be counted.
3. *Measure DNA contour length and protein-binding position* (Fig. 4B). In Image J, use the *Segmented Line Tool* (right-click on *Straight Line* to select *Segmented Line*) to measure the contour length and binding position. Left-click to begin the line and add vertices, along the length of the DNA; right-click to end the line. Select *Analyze* > *Measure* (or use shortcut "m") to add the current length measurement in units of pixels to the Results window. If images were captured as above, the conversion factor 1000 nm/512 pixels should be used to calculate length in appropriate units. Measure total contour length of the DNA molecule, as well as the length from the bound protein to the closest DNA end. Protein-binding position can be reported as percent from one end of the total DNA contour length, $P = (100 \times \text{length from DNA end to protein})/\text{total DNA contour length}$. Repeat for all labeled complexes.
4. *Measure DNA bend angle* (Fig. 4C). Use the *Angle Tool* in ImageJ to measure the DNA bend angle at the bound protein. Left-click to create the three points of the angle; these may be adjusted by dragging the points as desired. Place the middle point at the center of the bound protein, such that the angle measures the bend in the DNA immediately

adjacent on either side. Select *Analyze* > *Measure* (or use shortcut "m") to add the angle measurement (α) to the Results Window. DNA bend angles are typically reported as $\beta = 180 - \alpha$. Repeat for all labeled complexes.

5. *Measure complex volume* (Fig. 4A). Open the height channel from the flattened .001 file in Image SXM. Select *Analyze* > *Show Histogram* to display the distribution of heights in the image; record the mean (histogram peak) as the background for the image. Locate protein–DNA complexes (identified in step 2) and draw a line around the footprint with the eraser tool to demarcate it from naked DNA. Then, select *Options* > *Density Slice* to define the thresholds for analysis; set the upper threshold to its maximum and drag the lower threshold such that all particles (DNA and proteins) are highlighted with minimal background noise. To count particles, select *Analyze* > *Analyze Particles* and choose the following settings: Min Particle Size (pixels): 15; Max Particle Size (pixels): 999,999; Label Particles; Ignore Particles Touching Edge; Include Interior Holes; Reset Measurement Counter. Then, select *Analyze* > *Measure*, followed by *Analyze* > *Show Results*. This will open a new Results window, which can be copied into an Excel spreadsheet. "Mean" and "Area" are the average height of the particle and the area of its footprint, respectively; ensure that these values are reported in nm. Volumes (nm^3) are calculated as $V = (\text{mean} - \text{background}) \times \text{area}$.
6. Generate histograms of the binding positions (P), DNA bend angles (β), and complex volumes (V) using GraphPad Prism. Create a column table of the data and, in the Analysis toolbox, select *Analyze* > *Column analyses* > *Frequency distribution*. Adjust bin centers and widths as appropriate for the sample size (typically, the number of bins should be approximately $\sqrt{n}$). Under the *New graph* heading, select *Create a new graph of the results* and change the graph type to *XY graph, Histogram spikes*.
7. Viewing the histogram, select *Fit a curve with nonlinear regression* from the Analysis toolbox. Under the *Fit* tab, select *Gaussian* > *Gaussian*. This will generate a new page showing the parameters for the Gaussian fit to the histogram data. The best-fit values for mean and standard deviation can be used to describe the properties of the bound proteins.
8. Volume data can be further processed to infer molecular weights, and thus binding stoichiometry. Because AFM volumes are directly proportional to MW for most globular proteins (Ratcliff & Erie, 2001; Schneider, Larmer, Henderson, & Oberleithner, 1998), a standard curve can be generated and used for all experiments with the same probe type

and mode of data collection (Fig. 4E). Using the center of the fitted Gaussian as the mean volume, calculate the experimental MW to determine stoichiometry.

3.3.3 Notes

1. It may be useful to have the file open in the NanoScope Analysis software as well. The 3D view is helpful when identifying proteins on DNA, particularly in the case of smaller proteins.
2. In step 5, the minimum particle size may vary depending on the protein being studied and the threshold settings applied.
3. Sample data from different NER proteins are shown in Fig. 4. Wild-type Rad4–Rad23 binding to a 538-bp DNA substrate was analyzed for protein-binding position (Fig. 4F) and DNA bend angle (Fig. 4G) (Kong et al., 2016). A mutant form of UV-DDB (127 kDa) binding to a 538-bp DNA substrate was analyzed for protein volumes (Fig. 4D); the mean volume corresponds to a MW of 388.6 kDa, which suggests that the protein was bound as a dimer (Ghodke et al., 2014).

4. SINGLE-MOLECULE DNA TIGHTROPE ASSAY

To eliminate the need for constant flow and the potential of surface interactions, we have developed a unique optical platform, based on the ability to anchor both ends of a long DNA molecule on two nearby micron-sized poly-L-lysine-coated silica beads via electrostatic interaction, with the rest of the DNA suspended in between them, forming DNA tightropes (Fig. 5) (Kad et al., 2010). While the procedure involved does not offer the degree of precision and control afforded by the nanofabrication process used in constructing flow cells for DNA curtain assays (Gorman et al., 2007; Graneli, Yeykal, Robertson, & Greene, 2006; Lee et al., 2015; Sternberg, Redding, Jinek, Greene, & Doudna, 2014), its implementation is relatively straightforward. The DNA tightrope assay also elevates the DNA molecules, and therefore protein–DNA interactions, away from the coverslip, allowing complete access to elongated DNA in space and minimizing any potential adverse surface effects. To illuminate protein–DNA interactions taking place microns above the surface, a subcritical, oblique angle must be used to maximize the signal-to-noise ratio (Konopka & Bednarek, 2008; Tokunaga, Imamoto, & Sakata-Sogawa, 2008). Since its inception, we and others have utilized the DNA tightrope platform extensively to characterize proteins involved in prokaryotic and eukaryotic nucleotide and base excision repair,

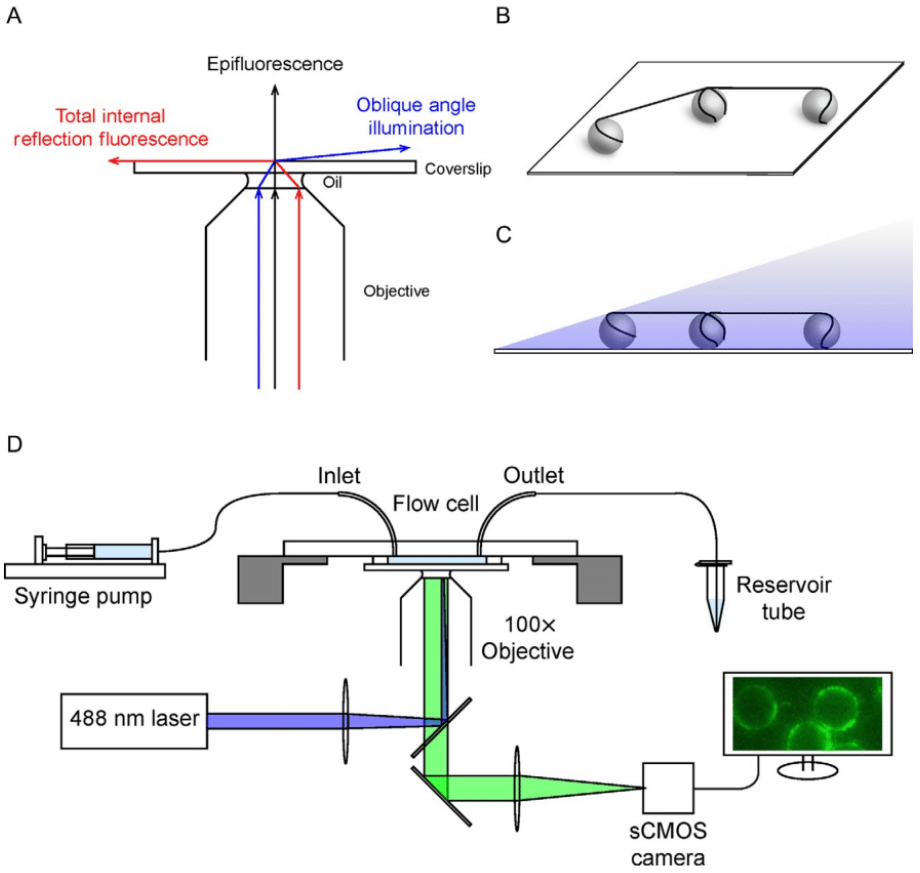

Fig. 5 Schematics of the DNA tightrope assay. (A) Schematic ray diagram of incident laser light paths for epifluorescence (*black*), total internal reflection fluorescence (TIRF) at the critical angle (*red*), and oblique angle illumination (*blue*). (B) Schematic of 5 μm poly-L-lysine-coated microspheres deposited on a glass coverslip with DNA tightropes suspended between them. (C) Schematic of DNA tightropes in the flow cell under oblique angle illumination. (D) Schematic of the experimental setup for the DNA tightrope assay. The flow cell is connected on the one end (*inlet*) to a syringe mounted on a syringe pump, while the other end (*outlet*) is connected to an Eppendorf tube reservoir. DNA tightropes in the flow cell are illuminated by a 488 nm laser (*blue*) under oblique angle through a 100× objective. Fluorescence signal (*green*) is imaged on a sCMOS camera connected to a computer.

as well as telomere shelterin complex components TRF1 and TRF2 (Dunn, Kad, Nelson, Warshaw, & Wallace, 2011; Ghodke et al., 2014; Hughes et al., 2013; Kong et al., 2016; Lin et al., 2014, 2016; Nelson et al., 2014). Due to the oblique angle illumination, the tightrope platform requires the use of Qdots to label proteins and provide sufficient fluorescence for visualization. These fluorescently stable and brilliant nanoparticles

allow continuous imaging at rates of 10–100 frames per second for collection periods of minutes without any photobleaching. Preparation of the flow cell begins with precoating clean coverslips with polyethylene glycol (Sections 4.1 and 4.2) and assembling predrilled microscope slides with inlet and outlet tubing (Section 4.3). Flow cells are constructed by attaching the coverslip to the slide assembly via a double-sided tape spacer (Section 4.5). Microspheres are simply flowed in such that they are distributed randomly but uniformly throughout the imaging area. Following deposition of the silica beads, tightropes are set up by continuously flowing DNA back and forth inside the flow cell for 40–60 min at the rate of 0.3 mL/min (Section 4.6). This step allows one end of the negatively charged DNA molecule to anchor to a positively charged bead, while the rest of the molecule is elongated by hydrodynamic force in the flow. With bead density optimized for length of DNA substrate used, the free end of the DNA molecule can attach to another bead in the vicinity. Proteins are visualized by Qdot labeling (Section 4.7), which is achieved either by conjugating a streptavidin-coated Qdot to a biotinylated antibody that recognizes the affinity tag on the protein (Ghodke et al., 2014; Kong et al., 2016) or through an antibody sandwich approach that utilizes a primary antibody against the affinity tag on the protein combined with a secondary antibody-coated Qdot (Kad et al., 2010; Wang, Tessmer, Croteau, Erie, & Van Houten, 2008). Data are collected, exported, and analyzed with a combination of software and scripts (Sections 4.8 and 4.9).

4.1 Cleaning Coverslips

4.1.1 Equipment

- Ultrasonic cleaning bath (Branson)
- Glass or plastic staining jars

4.1.2 Buffers and Reagents

- 20% Liquinox (ALCONOX)
- 100% acetone
- 100% ethanol
- 1 *M* potassium hydroxide (KOH)
- Coverslips (No. 1½, 24 × 40 mm, Corning)

4.1.3 Procedure

1. Load coverslips into staining jars and fill with 20% Liquinox detergent solution. Sonicate for 60 min.

2. Dump out detergent solution and rinse coverslips under deionized water until suds no longer form. Then fill staining jars with deionized water and sonicate for 5 min.
3. Replace deionized water in staining jars with acetone. Sonicate for 15 min.
4. Pour off acetone and rinse coverslips thoroughly under deionized water. Then fill staining jars with deionized water and sonicate for 5 min.
5. Replace deionized water in staining jars with 1 *M* KOH solution. Sonicate for 15 min.
6. Pour off and save KOH solution. Rinse coverslips thoroughly under deionized water. Then fill staining jars with 100% ethanol. Sonicate for 15 min.
7. Pour off ethanol and rinse coverslips thoroughly under deionized water. Then fill staining jars with 1 *M* KOH solution saved from the previous step. Sonicate for 15 min.
8. Pour off KOH solution and rinse coverslips thoroughly under deionized water. Then fill staining jars with deionized water and sonicate for 15 min.
9. Replace the deionized water in staining jars. Slides can be stored in water until they are to be used.

4.1.4 Notes

1. Do not allow coverslips to sit in 1 *M* KOH solution for prolonged time as they can be slowly etched by the solution.

4.2 PEGylation of Coverslips

4.2.1 Equipment

- Ultrasonic cleaning bath (Branson)
- Glass or plastic staining jars

4.2.2 Buffers and Reagents

- Aminosilane solution (for eight coverslips, 1.0 mL (3-aminopropyl) triethoxysilane, 2.5 mL glacial acetic acid, and 50 mL methanol, scale up if needed)
- 10 m*M* $NaHCO_3$, adjusted to pH ~8.5
- PEG solution (25 mg mPEG-succinimidyl valerate, MW 5000 (Laysan Bio) dissolved in 96 μL of $NaHCO_3$ solution)
- Compressed nitrogen gas

4.2.3 Procedure

1. Dry the cleaned coverslips completely with compressed nitrogen gas.
2. Let the coverslips sit in aminosilane solution for 20 min total. After 10 min, sonicate for 1 min and then sit for the remaining 9 min.
3. Pour off aminosilane solution and rinse the coverslips thoroughly under deionized water and dry with compressed nitrogen.
4. Prepare an empty tip box: fill it with deionized water up to a depth of ~1 cm and soak a piece of paper towel in the water.
5. Take a dry coverslip, mark the side that is *not* to be PEGylated with marker. Lay the coverslip marked-side-down on the tip rack. Take another coverslip, mark the side that is *not* to be PEGylated with marker. Set it aside, marked-side-down.
6. To create a coverslip "sandwich," deposit 20 μL of the PEG solution in the middle of the coverslip on the tip rack. Lay the other coverslip on top, marked-side-up. The liquid should spread out evenly between the two coverslips without forming any bubbles.
7. Repeat steps 5 and 6 for the remaining coverslips.
8. Shield the tip box from light with aluminum foil and place it in a dark place at room temperature overnight.
9. Disassemble the coverslip "sandwiches," place the coverslips in staining jars, and rinse them thoroughly under deionized water.
10. Blow dry coverslips with nitrogen gas and place them back on the tip rack, PEGylated side up (marked-side-down).
11. Cover the tip box with aluminum foil. PEGylated coverslips can be stored at 4°C for 2 weeks.

4.2.4 Notes

1. $NaHCO_3$ and PEG solutions can be prepared during step 2. PEG is especially light sensitive when in solution and should be protected from light.
2. After step 3, coverslips may be stored in methanol if the PEG solution is not ready.
3. Steps 6 through 10 should be carried out in a dark environment.
4. PEGylation (step 8) can be as short as 3 h, but overnight is preferred.

4.3 Assembly and Disassembly of Slides With Tubing (See Fig. 6)

4.3.1 Equipment

- Ultrasonic cleaning bath (Branson)
- Glass or plastic staining jars

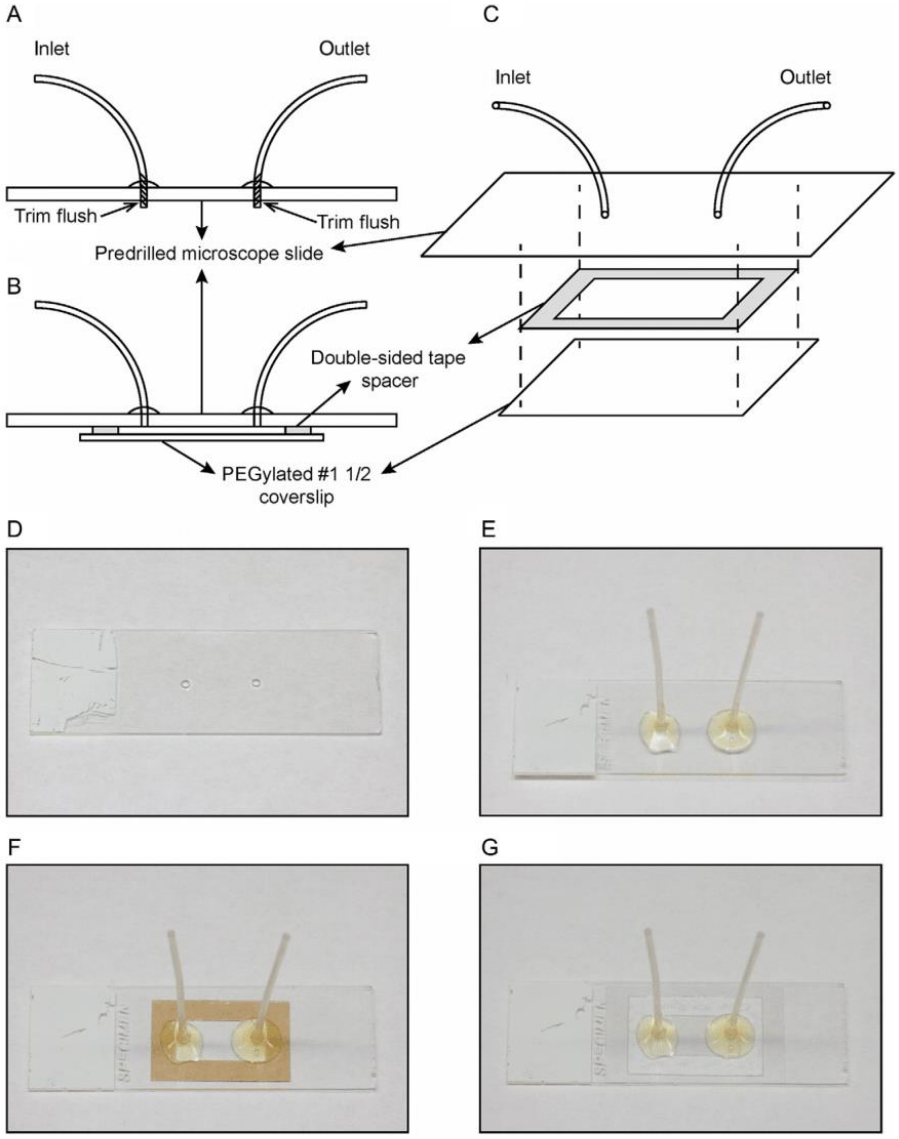

Fig. 6 Assembly of the flow cell. (A) Cross-sectional view of the predrilled microscope slide with inlet and outlet tubing attached. (B) Cross-sectional view of the assembled flow cell, where the double-sided tape spacer is sandwiched between the slide assembly and the PEGylated coverslip. (C) Exploded view of the flow cell assembly. (D) Microscope slide with predrilled holes. (E) Microscope slide with inlet and outlet tubing attached. (F) Microscope slide with inlet and outlet tubing and rectangular double-sided tape spacer (*brown*). (G) Complete flow cell assembly with glass coverslip attached to the microscope slide.

- Benchtop drill press and 1.25 mm diamond drill bit
- Extra fine grit sanding sponge (3M)

4.3.2 Buffers and Reagents

- Microscope slides (25 × 75 × 1 mm, Thermo Scientific)
- Teflon PFA tubing (1/16″ OD × 0.030″ ID, IDEX)
- Adhesive (BONDiT B-45TH, RELTEK)
- Slides cleaning solution (1 *M* HCl and 20% ethanol)
- 100% acetone
- 100% ethanol

4.3.3 Procedure

1. Drill two holes, 15–16 mm apart horizontally, in the center of a microscope slide (Fig. 6A and D). The precise distance between the holes is dependent on the desired size of the usable flow cell area.
2. Cut two pieces of the Teflon tubing to size, ~3 cm each. Rough up one end (~5 mm) of each piece of tubing with the sanding sponge for better adhesion.
3. Thread the roughed-up ends of the tubing through the holes in the slide. Apply adhesive around the base. Allow the ends to protrude ~1 mm from the other (bottom) side of the slide (Fig. 6A and E). This ensures that should some adhesive seeps through, it will not block the tubing.
4. Set the assembled slides aside at room temperature for at least 24–48 h to allow the adhesive to cure completely.
5. Drilled slides and Teflon tubing may be reused. For disassembly, submerge the flow cell (see below) in acetone for 1–2 days until it falls apart. Keep the slide and tubing and discard everything else. Remove any residual adhesive from the slide with a razor blade or KimWipe soaked in acetone.
6. In a staining jar, submerge used slides in acetone and sonicate for 1 h.
7. Discard acetone, rinse the slides thoroughly under deionized water, and fill the staining jar with the slides cleaning solution (1 *M* HCl and 20% ethanol). Sonicate for 1 h.
8. Discard the cleaning solution, rinse the slides thoroughly under deionized water, and fill the staining jar with 100% ethanol.
9. Wipe dry slides with KimWipes. Any remaining adhesive on the slides should be rubbed off with KimWipes and 100% ethanol.

4.3.4 Notes

1. It may be helpful to drill holes in the slide while it is submerged in water in order to help reduce the probability of slides cracking.

2. Some adhesives may cure faster (i.e., overnight) if the assembled slides are left in a 37°C incubator.

4.4 Preparation of Poly-L-Lysine-Coated Beads

4.4.1 Equipment

- Benchtop centrifuge
- Vertical rotators

4.4.2 Buffers and Reagents

- 5 μm silica microspheres (Polysciences)
- Poly-L-lysine powder (Waco Chemicals)

4.4.3 Procedure

1. Resuspend 100 μL of beads in 500 μL of ddH_2O. Centrifuge at 4°C for 4 min at 12,000 rpm.
2. Remove supernatant and resuspend beads in 400 μL of 2.5 mg/mL poly-L-lysine solution.
3. Rotate end to end at 4°C overnight on a vertical rotator.

4.4.4 Notes

1. 2.5 mg/mL poly-L-lysine solution is made in ddH_2O and can be stored at −20°C.
2. Poly-L-lysine-coated beads can be stored at 4°C.

4.5 Flow Cell Assembly (See Fig. 6)

4.5.1 Equipment

- Benchtop centrifuge
- Low-magnification light microscope

4.5.2 Buffers and Reagents

- Assembled predrilled slide with tubing
- Double-sided tape spacer
- PEGylated coverslip
- 200 μL gel-loading tips
- Blocking buffer (10 m*M* HEPES, pH 7.5, 50 m*M* NaCl, 1 mg/mL bovine serum albumin (Roche))
- Poly-L-lysine-coated silica beads

4.5.3 Procedure

1. Take a clean slide and use a razor blade to cut the protruding ends of tubing flush with the slide. Scrape back and forth to ensure that the bottom side of the slide is flat and smooth.
2. Cut out a double-sided tape spacer with a razor blade. Peel one side and paste it to the slide, using fingernail to firmly press the sticky tape (Fig. 6C and F).
3. Take one PEGylated coverslip from 4°C storage. Make sure that there is no excessive condensation or water on the treated (unmarked) surface. Hold the coverslip on its edges with fingers so that any condensation on the treated side evaporates quickly. Wipe the untreated (marked) surface dry with KimWipes.
4. Peel off the adhesive backing, make sure that the coverslip is completely dry, and place the PEGylated coverslip over the sticky tape spacer. Make sure the edges of the coverslip do not extend beyond those of the slide underneath it. Again, using the thumbnail, gently press around the outline of the spacer (Fig. 6B, C, and G).
5. With a 200-μL gel-loading tip, fill the flow cell with ~100 μL of the blocking buffer (10 m*M* HEPES, pH 7.5, 50 m*M* NaCl, 1 mg/mL BSA). Block the flow cell for 10 min.
6. After 10 min of initial blocking, examine the flow cell to ensure that no leakage has occurred, and then prepare the beads while blocking continues. First, vortex and resuspend the stock of beads in poly-L-lysine solution.
7. Add 13–15 μL of bead stock to 400 μL of ddH_2O. Resuspend again by vortexing and then centrifuge at 12,000 rpm for 4 min at 16°C. Carefully discard the supernatant without disturbing the pellet.
8. Repeat the washing step above with another 400 μL of ddH_2O. This time, after centrifugation, take out 300 μL of ddH_2O and then resuspend beads in the remaining ~110 μL.
9. Immediately after mixing, pipette ~110 μL of the suspension *slowly* into the flow cell with a gel-loading tip. Collect the bead flow-through and recirculate once if necessary.
10. Check the distribution and density of deposited beads in the flow cell with a low-magnification light microscope. Add more beads if necessary.
11. Allow the beads to settle for 10 min, and then flow 200 μL of ddH_2O through the flow cell to wash away any free beads.

4.5.4 Notes

1. Spacers can be prepared by folding a piece of double-sided tape (3M) onto itself to double the thickness and create two adhesive sides with removable backing. A nested-rectangle design pattern is then cut from the tape to make spacers. The outer rectangle should be slightly less than the size of the coverslip. The size of the inner rectangle corresponds to the usable flow cell area and should be large enough to encompass the predrilled holes in the microscope slide.
2. Poly-L-lysine-coated beads settle and clump together easily if left unperturbed. To ensure reproducible results, any pipetting should be done immediately after resuspension and vortexing. This is especially important in step 10.
3. It is useful to keep in mind the length of DNA tightropes to be used in the system when checking bead distribution and density. In order to determine whether enough beads have been deposited on the coverslip, compare the expected DNA tightrope length to inter-bead distances, which can be estimated based on known bead diameters.
4. The amounts of beads required may need to be further optimized with respect to the person carrying out this protocol.

4.6 Preparation of DNA Tightropes

4.6.1 Equipment

- Syringe pump (WPI)
- 5-mL glass syringe or plastic syringe (Hamilton)
- 21G hypodermic needle (BD)
- Teflon PFA tubing (1/16″ OD × 0.030″ ID, IDEX)
- Union assembly (0.020 through hole, for 1/16″ OD, IDEX)
- Flangeless ferrule (for 1/16″ OD, IDEX)

4.6.2 Buffers and Reagents

- Assembled flow cell
- 1 × TR buffer (20 m*M* HEPES, pH 7.5, 50 m*M* KCl, 3 m*M* $MgCl_2$)
- Long DNA substrate (λ-DNA or defined lesion substrates)

4.6.3 Procedure

1. Assemble the syringe, needle, tubing, and all fitting pieces. Wash the system by flowing ~5–6 mL of 1 × TR buffer through it. Leave ~1 mL in the syringe.
2. Set up and secure the syringe on the syringe pump. Set the flow rate to 0.3 mL/min, volume = 100 μL.
3. Connect the flow cell to the syringe by first pushing the TR buffer in the system through until the solution starts to drip from the female fitting piece that is to be connected. Quickly attach the fitting piece on the flow cell to that on the tubing.
4. Attach the outlet tubing to the other side of the flow cell and set up a predrilled Eppendorf tube as the reservoir. Add 500 μL of TR buffer to the reservoir tube and withdraw until there is only 1–2 μL left.
5. Thaw out DNA tightrope substrate, make up the volume to 100 μL with 1 × TR buffer. Vortex to resuspend well and spin down briefly to collect. Add DNA to reservoir tube and withdraw all.
6. Add 250 μL of 1 × TR buffer. Withdraw 100 μL to push the DNA from the outlet tubing into the flow cell. Set up the program to the continuous push–pull cycle (infusion followed by withdrawal) at the rate of 0.3 mL/min for a total volume of 100 μL in each direction.
7. Pause the syringe pump after 40–60 min of the continuous cycle.
8. If using ligated defined lesion damage arrays, wash the flow cell with 200 μL of 1 × high-salt TR buffer containing 1 *M* NaCl to remove DNA-bound ligase carried over from the ligation reaction. Then equilibrate the flow cell with 400 μL of protein binding buffer.

4.6.4 Notes

1. Introduction of air bubbles during step 3 is a common cause of failure. Attaching the flow cell to the tubing in a swift and smooth manner usually leads to better results. It is important to inspect the flow cell after step 3 for the presence of air bubbles. Small air bubbles trapped in the tubing that is attached to the flow cell can be backed out into the syringe and will not cause any issues downstream. A large column of air pushed into the flow cell will displace deposited beads, rendering the flow cell unusable.
2. The combination of the flow rate (0.3 mL/min) and time (40–60 min) of the continuous cycle employed to string up DNA tightropes has been shown to not overstretch DNA (Kad et al., 2010). Different

combinations can also be explored for potential effects on DNA tightrope conformations and protein binding.

4.7 Protein Conjugation

The use of oblique angle illumination for probing of protein–DNA interactions on tightropes that are suspended 5 μm above the surface requires the use of fluorescent probes that are exceptionally bright. Bioconjugated Qdots or QDs are commercially available and possess characteristics such as broad excitation spectrum and narrow size-dependent emission spectrum, as well as excellent brightness and photostability, all of which are highly beneficial to single-particle tracking (Bruchez, 2011). We have developed several approaches to label affinity purified proteins with Qdots for imaging on the tightrope platform, two of which are shown in Fig. 7. The first strategy takes advantage of the highly specific streptavidin–biotin interaction by conjugating streptavidin-coated Qdots with biotinylated antibodies against the affinity tag used in the purification of the protein of interest (Fig. 7A). Under certain circumstances, the placement of a relatively large Qdot close to the protein of interest may interfere with its ability to interact with other proteins or DNA. To prevent potential steric hindrance, we also developed the antibody sandwich approach, where a primary antibody against the affinity tag on the protein serves as the linker between a secondary antibody-coated Qdot and the affinity-tagged protein of interest (Wang et al., 2008) (Fig. 7B). Both approaches are straightforward to implement in one-color imaging of protein on λ-DNA or defined lesion substrates without biotin in the damage-containing oligonucleotide. However, to image more than one protein, it is essential to ensure that the Qdots on those proteins cannot exchange. The antibody sandwich approach can be easily adapted to this situation by using an orthogonal set of species of antibodies, i.e., goat-antimouse secondary antibody Qdots paired with mouse-anti-6xHis primary antibody, and goat-antirabbit secondary antibody Qdots paired with rabbit-anti-6xHis primary antibody. We have had great success at imaging two colors using this approach (Hughes et al., 2013), and depending on the optical setup with appropriate splitters and the number of protein tags, as many as six uniquely Qdot-labeled proteins could be feasibly imaged simultaneously.

4.7.1 Equipment

- Benchtop centrifuge

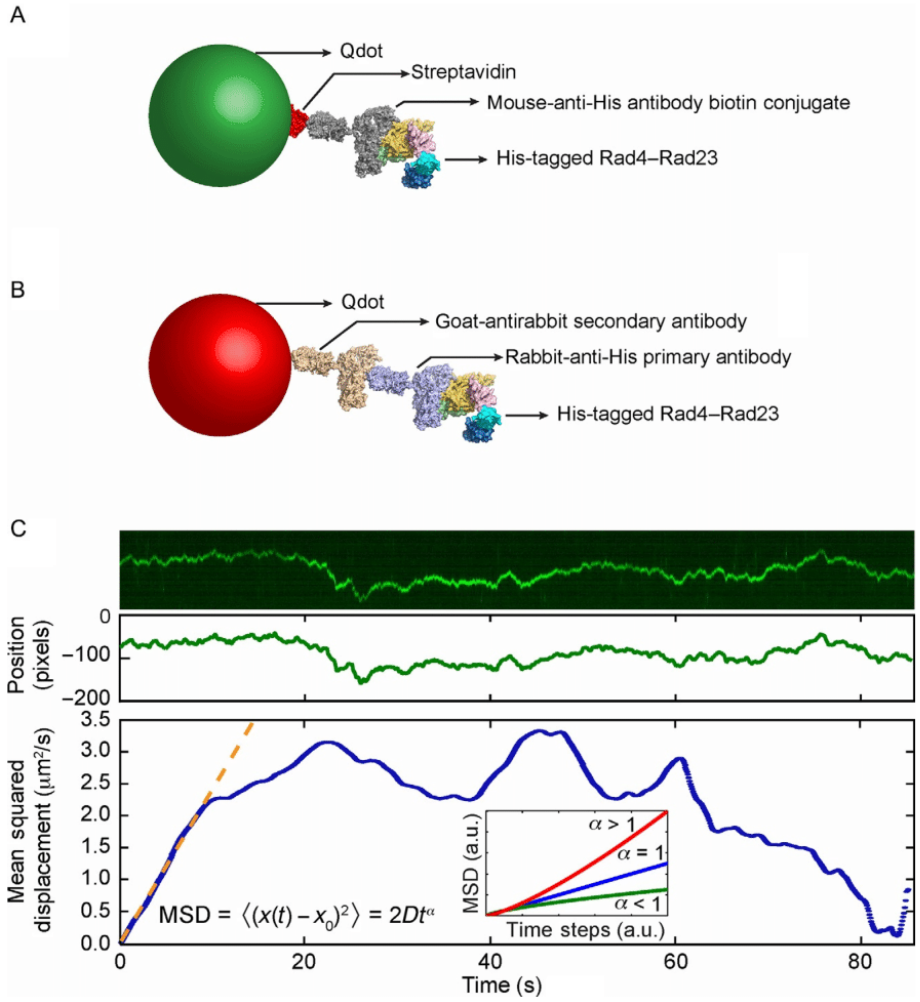

Fig. 7 Qdot conjugation strategies and data analysis. (A) Streptavidin (*red*)-coated quantum dot (*green*) is conjugated to His-tagged Rad4–Rad23 via the biotin-conjugated mouse-anti-His antibody (*gray*). (B) His-tagged Rad4–Rad is labeled by goat-antirabbit secondary antibody (*wheat*)-conjugated quantum dot (*red*) via a rabbit-anti-His primary antibody (*purple*). (C) *Top*: Representative kymograph of a diffusing particle. *Middle*: Plot of position, in the units of pixels (1 pixel = 46 nm), vs time, after fitting the light intensity profile at each time point in the kymograph with a one-dimensional Gaussian. *Bottom*: Plot of mean squared displacement (MSD), calculated from Gaussian-fitted positions, vs time steps. *Orange dashed line* is the result of fitting the initial portion of the MSD curve to the equation $\mathrm{MSD} = 2Dt^{\alpha}$. *Inset*: three types of one-dimensional diffusion characterized by different α values: superdiffusion (*red*), random diffusion (*blue*), and subdiffusion (*green*). Based on Movie 1 in the online version at http://dx.doi.org/10.1016/bs.mie.2017.03.027. *Adapted with permission from Kong, M., & Van Houten, B. (2016). Rad4 recognition-at-a-distance: Physical basis of conformation-specific anomalous diffusion of DNA repair proteins.* Progress in Biophysics and Molecular Biology. *http://dx.doi.org/10.1016/j.pbiomolbio.2016.12.004 (fig. 2C).*

4.7.2 Buffers and Reagents

- 1 μ*M* Qdot (streptavidin- or secondary antibody-conjugated, Invitrogen)
- 1 μ*M* biotin-conjugated anti-His antibody (Qiagen) or other anti-His primary antibody
- His-tagged protein of interest
- Protein storage buffer

4.7.3 Procedure

1. (a) For streptavidin-conjugated Qdots (SAQD): incubate 1 μL of 1 μ*M* of SAQD with 5 μL of 1 μ*M* of biotin-conjugated anti-His antibody (HisAb) so that the molar ratio of SAQD:HisAb is 1:5. Allow the binding reaction to proceed at 4°C for 1 h.
   (b) For secondary antibody-conjugated Qdots (IgGQD): incubate 1 μL of 1 μ*M* of anti-His primary antibody with 1 μL of 1 μ*M* of the His-tagged protein of interest and make up the volume with protein storage buffer to 5 μL. The molar ratio of protein:antibody is 1:1. Allow the binding reaction to proceed at 4°C for 1 h.
2. (a) For SAQD: incubate 1 μL of the mixture prepared in step 1(a) with 1 μL of 1/6 μ*M* of the protein of interest, such that the molar ratio of SAQD:HisAb:protein is 1:5:1 and the final concentration of the protein is ~83 n*M*. Allow the binding reaction to proceed at 4°C for 1 h.
   (b) For IgGQD: incubate 1 μL of the mixture prepared in step 1(b) with 1 μL of 1 μ*M* of IgGQD with the appropriate secondary antibody, such that the molar ratio of IgGQD:HisAb:protein is 5:1:1 and the final concentration of the protein is ~100 n*M*. Allow the binding reaction to proceed at 4°C for 1 h.

4.7.4 Notes

1. Depending on the stability of the protein of interest, conjugation steps may be carried out at room temperature to speed up the reaction.
2. Agarose gel-based electrophoretic mobility shift assays should be carried out with short DNA substrate, protein of interest, and each intermediate step of the Qdot conjugation protocol (i.e., protein with primary antibody, protein with primary antibody and Qdot) to ensure that DNA-binding activity is not lost due to conjugation of Qdot (Ghodke et al., 2014).

4.8 Data Collection

4.8.1 Equipment

- Benchtop centrifuge
- Inverted fluorescence microscope (Nikon Ti) with 100 × oil-based high-NA objective for TIRF-M, appropriate filter set for the wavelengths of Qdots used (optional), and high-speed sCMOS camera (Andor).
- Microscope user interface and image collection software (NIS-Elements Ar, Nikon)

4.8.2 Buffers and Reagents

- 1 × TR buffer (20 m*M* HEPES, pH 7.5, 50 m*M* KCl, 3 m*M* $MgCl_2$)
- 1 × high-salt TR buffer (20 m*M* HEPES, pH 7.5, 50 m*M* KCl, 3 m*M* $MgCl_2$, 1 *M* NaCl)
- Protein binding buffer
- QD-conjugated protein of interest
- Immersion oil

4.8.3 Procedure

1. Set up and secure the flow cell in the holder of the translational stage on the microscope, using immersion oil with appropriate 100 × objective lens. Focus the objective on the beads that have been deposited on the coverslip. Turn on any focus drift compensation if applicable (Perfect Focus System, Nikon).
2. Equilibrate the flow cell by passing through 4 volumes (400 μL) of protein binding buffer from the reservoir tube.
3. Dilute 1 μL of QD-conjugated protein of interest in 100 μL of protein binding buffer. Pipette the diluted protein solution into the reservoir tube and withdraw all. The final concentration of QD-labeled protein is ∼1 n*M*.
4. Pipette 100 μL of protein binding buffer into the reservoir tube and withdraw all, such that the protein solution is pushed into the flow cell.
5. Turn on the excitation laser and find the critical TIRF angle where QD fluorescence can just begin to be observed. Then increase the angle slightly to optimize for signal-to-noise ratio.
6. Move the translational stage to look for binding events in the live-view window.
7. After locating a region of interest, set up recording frame rate (∼10 fps), time (∼5–15 min), and file directory. If more than one Qdot

wavelength is used, configure the emission filter as needed. Record the time series.

8. Refresh QD-labeled proteins at least every 2 h, depending on the stability of protein while under the microscope.

4.8.4 Notes

1. When using tandem-ligated plasmid DNA substrates, wash the flow cell with 200 μL of 1 × high-salt TR buffer containing 1 *M* NaCl prior to equilibration with protein-binding buffer to remove DNA-bound ligase carried over from the ligation reaction.
2. It is important to perform negative controls with QD–HisAb complexes only, in the absence of protein conjugation, to confirm that they do not stick to DNA in a nonspecific manner.
3. If DNA binding is rare, consider increasing the concentration of QD-labeled protein in the flow cell. The empirical maximum concentration of fluorescent Qdots, including both free and protein-conjugated, is ∼10 n*M*. Background fluorescence from freely diffusing Qdot could become overwhelming above this limit.

4.9 Data Analysis

4.9.1 Equipment

- Image processing software (NIS-Elements Ar or NIS-Elements Viewer, Nikon)
- ImageJ (https://imagej.nih.gov/ij/, NIH)
- Data processing and fitting software (Matlab, MathWorks)

4.9.2 Procedure

1. Convert manufacturer-specific proprietary image stack file format (.nd2, Nikon) to a time series of individual TIFF files. Separate the channels if multiple Qdot emission wavelengths are used.
2. Import the time series of TIFF files as an *Image sequence* in ImageJ. Save the image stack in ImageJ as a single TIFF file. For an example of a time series, see Movie 1 in the online version at http://dx.doi.org/10.1016/bs.mie.2017.03.027.
3. Using the *Straight line* tool in the tool bar, trace the linear trajectory of one-dimensional diffusion of one QD-labeled particle. Ensure that the length of the line covers the entire range of motion.
4. Press the "/" key or go to *Image* > *Stacks* > *Reslice*. In the *Reslice* window that pops up, check the box *Rotate 90 degrees*. Click *OK* to generate a

kymograph that displays the particle position (on the vertical axis) over time (on the horizontal axis). Save the kymograph as a TIFF file.

5. Fit the fluorescence intensity in the kymograph with a one-dimensional Gaussian fitting algorithm in ImageJ. Save the Gaussian-fitted peak positions (Fig. 7C).
6. In Matlab, or other appropriate data processing software, import the Gaussian-fitted peak positions and calculate the one-dimensional mean square displacement (MSD) as a function of time steps

$$\mathrm{MSD}(n\Delta t) = \frac{1}{N-n}\sum_{i=1}^{N-n}(x_{i+n}-x_i)^2,$$

where N is the total number of frames in the time series, n is the number of frames for different time steps, x_i is the Gaussian-fitted peak position in the ith frame, and Δt is the unit time step between consecutive frames, i.e., the inverse of the frame rate.

7. Extract diffusion coefficient D and anomalous diffusion exponent α from the MSD by fitting the equation

$$\mathrm{MSD} = 2Dt^{\alpha}$$

Begin the fitting process by using all available data points in the MSD curve. In each round of fitting, reduce the number of data points used by one, taken from the end of the MSD curve, until desired goodness of fit is achieved (Fig. 7C). For an example of Matlab script, see MSD_main.m in the supplementary file (http://dx.doi.org.10.1016/bs.mie.2017.03.027).

4.9.3 Notes

1. It is important to first establish the systematic noise level of the platform, in terms of the one-dimensional diffusion coefficient value of stably bound nonmotile Qdots on the tightrope.
2. By analyzing the component of diffusive motion that is along the direction of the tightrope (longitudinal), an implicit assumption is made that particle motion perpendicular to the tightrope (transverse) is at the background noise level. This assumption can be verified by observing that the particle of interest, motile or nonmotile, does not exhibit any kind of "wobble" on the tightrope, whose direction is in general parallel to that of the hydrodynamic flow expected in the flow cell. Quantitatively, two-dimensional tracking of the particle can be employed to determine its x and y positions. In practice, particles that exhibit any "wobble" on

the tightrope should be excluded from further analysis as the behavior indicates that the tightrope itself is not anchored properly on beads or has structural defects.

3. In the case of multiple binding events on one DNA tightrope, it is beneficial to extract the kymographs of all particles on the tightrope, motile and nonmotile, by drawing one straight line through all the particles. Kymographs of individual particles can be cropped out and analyzed independently.
4. An ImageJ script for one-dimensional Gaussian fitting is available for download at http://kadlab.mechanicsanddynamics.com/images/Downloads/Gaussian_Fit.txt (Kad et al., 2010).
5. Resolution of the system can be characterized by the positional accuracy (Thompson, Larson, & Webb, 2002) and localization precision (Arnspang, Brewer, & Lagerholm, 2012). Calculations of these quantities relevant to the tightrope platform have been detailed elsewhere (Ghodke et al., 2014).

5. CONCLUSIONS

In summary, we have established a complete laboratory workflow from bulk biochemistry to single-molecule biophysics. The experimental platform detailed in this chapter is well suited for characterization of not just proteins involved in NER, but protein–DNA interactions in general. Specifically, the DNA tightrope assay is straightforward to implement and its versatility allows the technique to be applied to investigate repair pathways such as base excision repair and mismatch repair, as well as the target search process of telomere shelterin complex proteins (Lin et al., 2014). Tightropes have also been constructed from actin filaments to study the cooperative activation of thin filaments (Desai, Geeves, & Kad, 2015). In addition to the dynamic and transient behavior observable on DNA tightropes, the use of AFM allows independent snapshot measurements of specific and nonspecific binding in the absence of any labeling fluorescent probes and visualization of any mechanical changes in DNA conformation that can be induced through protein binding. Both complementary techniques benefit greatly from the utilization of defined lesion substrates such that specific binding events can be more readily differentiated from nonspecific ones. In the future, the challenges ahead lie in the development of incorporating nucleosomes (Lee & Greene, 2011; Visnapuu & Greene, 2009) in the defined lesion damage arrays, as well as complete reconstitutions of repair

pathways at the single-molecule level with efficient real-time multicolor imaging capabilities.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/bs.mie.2017.03.027.

ACKNOWLEDGMENTS

This work was made possible through funding from the National Institutes of Health 5R01ES019566 to B.V.H., and 2P30CA047904 to University of Pittsburgh Cancer Institute.

REFERENCES

Arnspang, E. C., Brewer, J. R., & Lagerholm, B. C. (2012). Multi-color single particle tracking with quantum dots. *PLoS One*, 7(11), e48521. http://dx.doi.org/10.1371/journal.pone.0048521

Bruchez, M. P. (2011). Quantum dots find their stride in single molecule tracking. *Current Opinion in Chemical Biology*, *15*(6), 775–780. http://dx.doi.org/10.1016/j.cbpa.2011.10.011.

Croteau, D. L., DellaVecchia, M. J., Perera, L., & Van Houten, B. (2008). Cooperative damage recognition by UvrA and UvrB: Identification of UvrA residues that mediate DNA binding. *DNA Repair (Amst)*, 7(3), 392–404. http://dx.doi.org/10.1016/j.dnarep.2007.11.013.

Croteau, D. L., DellaVecchia, M. J., Wang, H., Bienstock, R. J., Melton, M. A., & Van Houten, B. (2006). The C-terminal zinc finger of UvrA does not bind DNA directly but regulates damage-specific DNA binding. *The Journal of Biological Chemistry*, *281*(36), 26370–26381. http://dx.doi.org/10.1074/jbc.M603093200.

Desai, R., Geeves, M. A., & Kad, N. M. (2015). Using fluorescent myosin to directly visualize cooperative activation of thin filaments. *The Journal of Biological Chemistry*, *290*(4), 1915–1925. http://dx.doi.org/10.1074/jbc.M114.609743.

Dunn, A. R., Kad, N. M., Nelson, S. R., Warshaw, D. M., & Wallace, S. S. (2011). Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. *Nucleic Acids Research*, *39*(17), 7487–7498. http://dx.doi.org/10.1093/nar/gkr459.

Furda, A. M., Bess, A. S., Meyer, J. N., & Van Houten, B. (2012). Analysis of DNA damage and repair in nuclear and mitochondrial DNA of animal cells using quantitative PCR. *Methods in Molecular Biology*, *920*, 111–132. http://dx.doi.org/10.1007/978-1-61779-998-3_9.

Geng, H., Du, C., Chen, S., Salerno, V., Manfredi, C., & Hsieh, P. (2011). In vitro studies of DNA mismatch repair proteins. *Analytical Biochemistry*, *413*(2), 179–184. http://dx.doi.org/10.1016/j.ab.2011.02.017.

Ghodke, H., Wang, H., Hsieh, C. L., Woldemeskel, S., Watkins, S. C., Rapic-Otrin, V., et al. (2014). Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(18), E1862–E1871. http://dx.doi.org/10.1073/pnas.1323856111.

Gorman, J., Chowdhury, A., Surtees, J. A., Shimada, J., Reichman, D. R., Alani, E., et al. (2007). Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Molecular Cell*, *28*(3), 359–370. http://dx.doi.org/10.1016/j.molcel.2007.09.008.

Graneli, A., Yeykal, C. C., Robertson, R. B., & Greene, E. C. (2006). Long-distance lateral diffusion of human Rad51 on double-stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1221–1226. http://dx.doi.org/10.1073/pnas.0508366103.

Hansma, H. G., & Laney, D. E. (1996). DNA binding to mica correlates with cationic radius: Assay by atomic force microscopy. *Biophysical Journal*, *70*(4), 1933–1939. http://dx.doi.org/10.1016/S0006-3495(96)79757-6.

Hughes, C. D., Wang, H., Ghodke, H., Simons, M., Towheed, A., Peng, Y., et al. (2013). Real-time single-molecule imaging reveals a direct interaction between UvrC and UvrB on DNA tightropes. *Nucleic Acids Research*, *41*(9), 4901–4912. http://dx.doi.org/10.1093/nar/gkt177.

Kad, N. M., Wang, H., Kennedy, G. G., Warshaw, D. M., & Van Houten, B. (2010). Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by single- molecule imaging of quantum-dot-labeled proteins. *Molecular Cell*, *37*(5), 702–713. http://dx.doi.org/10.1016/j.molcel.2010.02.003.

Kong, M., Liu, L., Chen, X., Driscoll, K. I., Mao, P., Bohm, S., et al. (2016). Single-molecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process. *Molecular Cell*, *64*(2), 376–387. http://dx.doi.org/10.1016/j.molcel.2016.09.005.

Kong, M., & Van Houten, B. (2016). Rad4 recognition-at-a-distance: Physical basis of conformation-specific anomalous diffusion of DNA repair proteins. *Progress in Biophysics and Molecular Biology*. http://dx.doi.org/10.1016/j.pbiomolbio.2016.12.004

Konopka, C. A., & Bednarek, S. Y. (2008). Variable-angle epifluorescence microscopy: A new way to look at protein dynamics in the plant cell cortex. *The Plant Journal*, *53*(1), 186–196. http://dx.doi.org/10.1111/j.1365-313X.2007.03306.x.

Lee, J. Y., & Greene, E. C. (2011). Assembly of recombinant nucleosomes on nanofabricated DNA curtains for single-molecule imaging. *Methods in Molecular Biology*, *778*, 243–258. http://dx.doi.org/10.1007/978-1-61779-261-8_16.

Lee, J. Y., Terakawa, T., Qi, Z., Steinfeld, J. B., Redding, S., Kwon, Y., et al. (2015). DNA RECOMBINATION. Base triplet stepping by the Rad51/RecA family of recombinases. *Science*, *349*(6251), 977–981. http://dx.doi.org/10.1126/science.aab2666.

Lin, J., Countryman, P., Buncher, N., Kaur, P., Longjiang, E., Zhang, Y., et al. (2014). TRF1 and TRF2 use different mechanisms to find telomeric DNA but share a novel mechanism to search for protein partners at telomeres. *Nucleic Acids Research*, *42*(4), 2493–2504. http://dx.doi.org/10.1093/nar/gkt1132.

Lin, J., Countryman, P., Chen, H., Pan, H., Fan, Y., Jiang, Y., et al. (2016). Functional interplay between SA1 and TRF1 in telomeric DNA binding and DNA-DNA pairing. *Nucleic Acids Research*, *44*(13), 6363–6376. http://dx.doi.org/10.1093/nar/gkw518.

Meyer, J. N., Boyd, W. A., Azzam, G. A., Haugen, A. C., Freedman, J. H., & Van Houten, B. (2007). Decline of nucleotide excision repair capacity in aging Caenorhabditis elegans. *Genome Biology*, *8*(5), R70. http://dx.doi.org/10.1186/gb-2007-8-5-r70.

Nelson, S. R., Dunn, A. R., Kathe, S. D., Warshaw, D. M., & Wallace, S. S. (2014). Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(20), E2091–E2099. http://dx.doi.org/10.1073/pnas.1400386111.

Ratcliff, G. C., & Erie, D. A. (2001). A novel single-molecule study to determine protein–protein association constants. *Journal of the American Chemical Society*, *123*(24), 5632–5635.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: A laboratory manual* (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Schneider, S. W., Larmer, J., Henderson, R. M., & Oberleithner, H. (1998). Molecular weights of individual proteins correlate with molecular volumes measured by atomic

force microscopy. *Pflügers Archiv*, *435*(3), 362–367. http://dx.doi.org/10.1007/s004240050524.

Springall, L., Inchingolo, A. V., & Kad, N. M. (2016). DNA-protein interactions studied directly using single molecule fluorescence imaging of quantum dot tagged proteins moving on DNA tightropes. *Methods in Molecular Biology*, *1431*, 141–150. http://dx.doi.org/10.1007/978-1-4939-3631-1_11.

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, *507*(7490), 62–67. http://dx.doi.org/10.1038/nature13011.

Tafvizi, A., Huang, F., Fersht, A. R., Mirny, L. A., & van Oijen, A. M. (2011). A single-molecule characterization of p53 search on DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(2), 563–568. http://dx.doi.org/10.1073/pnas.1016020107.

Thompson, R. E., Larson, D. R., & Webb, W. W. (2002). Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, *82*(5), 2775–2783. http://dx.doi.org/10.1016/S0006-3495(02)75618-X.

Tokunaga, M., Imamoto, N., & Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature Methods*, *5*(2), 159–161. http://dx.doi.org/10.1038/nmeth1171.

Vesenka, J., Guthold, M., Tang, C. L., Keller, D., Delaine, E., & Bustamante, C. (1992). Substrate preparation for reliable imaging of DNA molecules with the scanning force microscope. *Ultramicroscopy*, *42–44*(Pt B), 1243–1249.

Visnapuu, M. L., & Greene, E. C. (2009). Single-molecule imaging of DNA curtains reveals intrinsic energy landscapes for nucleosome deposition. *Nature Structural & Molecular Biology*, *16*(10), 1056–1062. http://dx.doi.org/10.1038/nsmb.1655.

Wang, H., DellaVecchia, M. J., Skorvaga, M., Croteau, D. L., Erie, D. A., & Van Houten, B. (2006). UvrB domain 4, an autoinhibitory gate for regulation of DNA binding and ATPase activity. *The Journal of Biological Chemistry*, *281*(22), 15227–15237. http://dx.doi.org/10.1074/jbc.M601476200.

Wang, H., Tessmer, I., Croteau, D. L., Erie, D. A., & Van Houten, B. (2008). Functional characterization and atomic force microscopy of a DNA repair protein conjugated to a quantum dot. *Nano Letters*, *8*(6), 1631–1637. http://dx.doi.org/10.1021/nl0803161.

Yeh, J. I., Levine, A. S., Du, S., Chinte, U., Ghodke, H., Wang, H., et al. (2012). Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(41), E2737–E2746. http://dx.doi.org/10.1073/pnas.1110067109.

Bibliography

1 Berg, O. G., Winter, R. B. & von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929-6948 (1981).

2 Winter, R. B., Berg, O. G. & von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor--operator interaction: kinetic measurements and conclusions. *Biochemistry* **20**, 6961-6977 (1981).

3 von Hippel, P. H. & Berg, O. G. On the specificity of DNA-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 1608-1612, doi:10.1073/pnas.83.6.1608 (1986).

4 von Hippel, P. H. & Berg, O. G. Facilitated target location in biological systems. *J Biol Chem* **264**, 675-678 (1989).

5 Slutsky, M. & Mirny, L. A. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophysical journal* **87**, 4021-4035, doi:10.1529/biophysj.104.050765 (2004).

6 Wang, F. *et al.* The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion. *Nat Struct Mol Biol* **20**, 174-181, doi:10.1038/nsmb.2472 (2013).

7 Kalisky, T., Dekel, E. & Alon, U. Cost-benefit theory and optimal design of gene regulation functions. *Phys Biol* **4**, 229-245, doi:10.1088/1478-3975/4/4/001 (2007).

8 Araujo, S. J., Nigg, E. A. & Wood, R. D. Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. *Mol Cell Biol* **21**, 2281-2291, doi:10.1128/MCB.21.7.2281-2291.2001 (2001).

9 van der Spek, P. J. *et al.* XPC and human homologs of RAD23: intracellular localization and relationship to other nucleotide excision repair complexes. *Nucleic Acids Res* **24**, 2551-2559, doi:10.1093/nar/24.13.2551 (1996).

10 Adam, G. & Delbrück, M. Reduction of dimensionality in biological diffusion processes. *Structural chemistry and molecular biology* **198**, 198-215 (1968).

11 Riggs, A. D., Bourgeois, S. & Cohn, M. The lac repressor-operator interaction. 3. Kinetic studies. *J Mol Biol* **53**, 401-417, doi:10.1016/0022-2836(70)90074-4 (1970).

12 Richter, P. H. & Eigen, M. Diffusion controlled reaction rates in spheroidal geometry. Application to repressor--operator association and membrane bound enzymes. *Biophys Chem* **2**, 255-263, doi:10.1016/0301-4622(74)80050-5 (1974).

13 Slutsky, M., Kardar, M. & Mirny, L. A. Diffusion in correlated random potentials, with applications to DNA. *Phys Rev E Stat Nonlin Soft Matter Phys* **69**, 061903, doi:10.1103/PhysRevE.69.061903 (2004).

14 Mirny, L. *et al.* How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *Journal of Physics A: Mathematical and Theoretical* **42**, 434013 (2009).

15 Hu, T. & Shklovskii, B. I. How does a protein search for the specific site on DNA: The role of disorder. *Phys Rev E Stat Nonlin Soft Matter Phys* **74**, 021903, doi:10.1103/PhysRevE.74.021903 (2006).

16 Tafvizi, A., Huang, F., Fersht, A. R., Mirny, L. A. & van Oijen, A. M. A single-molecule characterization of p53 search on DNA. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 563-568, doi:10.1073/pnas.1016020107 (2011).

17 Leith, J. S. *et al.* Sequence-dependent sliding kinetics of p53. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16552-16557, doi:10.1073/pnas.1120452109 (2012).

18 Gorman, J. *et al.* Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Mol Cell* **28**, 359-370, doi:10.1016/j.molcel.2007.09.008 (2007).

19 Buechner, C. N., Maiti, A., Drohat, A. C. & Tessmer, I. Lesion search and recognition by thymine DNA glycosylase revealed by single molecule imaging. *Nucleic Acids Research* **43**, 2716-2729, doi:10.1093/nar/gkv139 (2015).

20 DeRocco, V. C., Sass, L. E., Qiu, R., Weninger, K. R. & Erie, D. A. Dynamics of MutS-mismatched DNA complexes are predictive of their repair phenotypes. *Biochemistry* **53**, 2043-2052, doi:10.1021/bi401429b (2014).

21 Ghodke, H. *et al.* Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E1862-1871, doi:10.1073/pnas.1323856111 (2014).

22 Kong, M. *et al.* Single-Molecule Imaging Reveals that Rad4 Employs a Dynamic DNA Damage Recognition Process. *Mol Cell* **64**, 376-387, doi:10.1016/j.molcel.2016.09.005 (2016).

23 Liu, L. *et al.* PARP1 changes from three-dimensional DNA damage searching to one-dimensional diffusion after auto-PARylation or in the presence of APE1. *Nucleic Acids Res* **45**, 12834-12847, doi:10.1093/nar/gkx1047 (2017).

24 Lindahl, T. & Barnes, D. E. Repair of endogenous DNA damage. *Cold Spring Harb Symp Quant Biol* **65**, 127-133 (2000).

25 Friedberg, E. C. *et al.* DNA repair: from molecular mechanism to human disease. *DNA Repair (Amst)* **5**, 986-996 (2006).

26 Mitchell, D. L. & Nairn, R. S. The biology of the (6-4) photoproduct. *Photochem Photobiol* **49**, 805-819, doi:10.1111/j.1751-1097.1989.tb05578.x (1989).

27 Kim, J. K., Patel, D. & Choi, B. S. Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity. *Photochem Photobiol* **62**, 44-50, doi:10.1111/j.1751-1097.1995.tb05236.x (1995).

28 Fichtinger-Schepman, A. M., van der Veer, J. L., den Hartog, J. H., Lohman, P. H. & Reedijk, J. Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. *Biochemistry* **24**, 707-713, doi:10.1021/bi00324a025 (1985).

29 Reardon, J. T., Vaisman, A., Chaney, S. G. & Sancar, A. Efficient nucleotide excision repair of cisplatin, oxaliplatin, and Bis-aceto-ammine-dichloro-cyclohexylamine-platinum(IV) (JM216) platinum intrastrand DNA diadducts. *Cancer Res* **59**, 3968-3971 (1999).

30 Gunz, D., Hess, M. T. & Naegeli, H. Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *J Biol Chem* **271**, 25089-25098, doi:10.1074/jbc.271.41.25089 (1996).

31 Hess, M. T., Gunz, D. & Naegeli, H. A repair competition assay to assess recognition by human nucleotide excision repair. *Nucleic Acids Res* **24**, 824-828, doi:10.1093/nar/24.5.824 (1996).

32 Mu, H. *et al.* Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies. *Nucleic Acids Res* **40**, 9675-9690, doi:10.1093/nar/gks788 (2012).

33 Yeo, J. E., Khoo, A., Fagbemi, A. F. & Scharer, O. D. The efficiencies of damage recognition and excision correlate with duplex destabilization induced by acetylaminofluorene adducts in human nucleotide excision repair. *Chem Res Toxicol* **25**, 2462-2468, doi:10.1021/tx3003033 (2012).

34 Van Houten, B. & Kong, M. in *Encyclopedia of Cell Biology* Vol. 1 (eds Ralph A. Bradshaw & Philip D. Stahl) 435-441 (Academic Press, 2016).

35 Sugasawa, K. Molecular mechanisms of DNA damage recognition for mammalian nucleotide excision repair. *DNA Repair (Amst)* **44**, 110-117, doi:10.1016/j.dnarep.2016.05.015 (2016).

36 Scharer, O. D. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol* **5**, a012609, doi:10.1101/cshperspect.a012609 (2013).

37 Epanchintsev, A. *et al.* Cockayne's Syndrome A and B Proteins Regulate Transcription Arrest after Genotoxic Stress by Promoting ATF3 Degradation. *Mol Cell* **68**, 1054-1066 e1056, doi:10.1016/j.molcel.2017.11.009 (2017).

38 Guerrero-Santoro, J. *et al.* The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A. *Cancer Res* **68**, 5014-5022, doi:10.1158/0008-5472.CAN-07-6162 (2008).

39 Luijsterburg, M. S. *et al.* DDB2 promotes chromatin decondensation at UV-induced DNA damage. *J Cell Biol* **197**, 267-281, doi:10.1083/jcb.201106074 (2012).

40 Matsumoto, S. *et al.* DNA damage detection in nucleosomes involves DNA register shifting. *Nature* **571**, 79-84, doi:10.1038/s41586-019-1259-3 (2019).

41 Kusakabe, M. *et al.* Mechanism and regulation of DNA damage recognition in nucleotide excision repair. *Genes Environ* **41**, 2, doi:10.1186/s41021-019-0119-6 (2019).

42 Dip, R., Camenisch, U. & Naegeli, H. Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. *DNA Repair (Amst)* **3**, 1409-1423, doi:10.1016/j.dnarep.2004.05.005 (2004).

43 Egly, J. M. & Coin, F. A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair (Amst)* **10**, 714-721, doi:10.1016/j.dnarep.2011.04.021 (2011).

44 Hohl, M. *et al.* Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity. *Nucleic Acids Res* **35**, 3053-3063, doi:10.1093/nar/gkm092 (2007).

45 Staresincic, L. *et al.* Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J* **28**, 1111-1120, doi:10.1038/emboj.2009.49 (2009).

46 Wakasugi, M. & Sancar, A. Order of assembly of human DNA repair excision nuclease. *J Biol Chem* **274**, 18759-18768 (1999).

47 Volker, M. *et al.* Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* **8**, 213-224 (2001).

48 Nocentini, S., Coin, F., Saijo, M., Tanaka, K. & Egly, J. M. DNA damage recognition by XPA protein promotes efficient recruitment of transcription factor II H. *J Biol Chem* **272**, 22991-22994 (1997).

49 Li, L., Elledge, S. J., Peterson, C. A., Bales, E. S. & Legerski, R. J. Specific association between the human DNA repair proteins XPA and ERCC1. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5012-5016 (1994).

50 Buchko, G. W. *et al.* Interactions of human nucleotide excision repair protein XPA with DNA and RPA70 Delta C327: chemical shift mapping and 15N NMR relaxation studies. *Biochemistry* **38**, 15116-15128 (1999).

51 Mer, G. *et al.* Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell* **103**, 449-456 (2000).

52 Wakasugi, M. *et al.* Damaged DNA-binding protein DDB stimulates the excision of cyclobutane pyrimidine dimers in vitro in concert with XPA and replication protein A. *J Biol Chem* **276**, 15434-15440, doi:10.1074/jbc.M011177200 (2001).

53 Gilljam, K. M., Muller, R., Liabakk, N. B. & Otterlei, M. Nucleotide excision repair is associated with the replisome and its efficiency depends on a direct interaction between XPA and PCNA. *PLoS One* **7**, e49199, doi:10.1371/journal.pone.0049199 (2012).

54 Sugitani, N., Sivley, R. M., Perry, K. E., Capra, J. A. & Chazin, W. J. XPA: A key scaffold for human nucleotide excision repair. *DNA Repair (Amst)* **44**, 123-135, doi:10.1016/j.dnarep.2016.05.018 (2016).

55 Li, L., Lu, X., Peterson, C. A. & Legerski, R. J. An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Mol Cell Biol* **15**, 5396-5402 (1995).

56 Saijo, M., Takedachi, A. & Tanaka, K. Nucleotide excision repair by mutant xeroderma pigmentosum group A (XPA) proteins with deficiency in interaction with RPA. *J Biol Chem* **286**, 5476-5483, doi:10.1074/jbc.M110.172916 (2011).

57 Matsuda, T. *et al.* DNA repair protein XPA binds replication protein A (RPA). *J Biol Chem* **270**, 4152-4157 (1995).

58 Camenisch, U. & Nageli, H. XPA gene, its product and biological roles. *Adv Exp Med Biol* **637**, 28-38 (2008).

59 Lin, J. J. & Sancar, A. (A)BC excinuclease: the Escherichia coli nucleotide excision repair enzyme. *Mol Microbiol* **6**, 2219-2224 (1992).

60 Reardon, J. T. & Sancar, A. Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes Dev* **17**, 2539-2551, doi:10.1101/gad.1131003 (2003).

61 Sugasawa, K., Akagi, J., Nishi, R., Iwai, S. & Hanaoka, F. 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, doi:10.1016/j.molcel.2009.09.035 (2009).

62 Li, C. L. *et al.* Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Mol Cell* **59**, 1025-1034, doi:10.1016/j.molcel.2015.08.012 (2015).

63 Bukowska, B. & Karwowski, B. T. Actual state of knowledge in the field of diseases related with defective nucleotide excision repair. *Life Sci* **195**, 6-18, doi:10.1016/j.lfs.2017.12.035 (2018).

64 Jaarsma, D., van der Pluijm, I., van der Horst, G. T. & Hoeijmakers, J. H. Cockayne syndrome pathogenesis: lessons from mouse models. *Mech Ageing Dev* **134**, 180-195, doi:10.1016/j.mad.2013.04.003 (2013).

65 Faghri, S., Tamura, D., Kraemer, K. H. & Digiovanna, J. J. Trichothiodystrophy: a systematic review of 112 published cases characterises a wide spectrum of clinical manifestations. *J Med Genet* **45**, 609-621, doi:10.1136/jmg.2008.058743 (2008).

66 Iakoucheva, L. M. *et al.* Aberrant mobility phenomena of the DNA repair protein XPA. *Protein Sci* **10**, 1353-1362, doi:10.1110/ps.ps.40101 (2001).

67 Iakoucheva, L. M. *et al.* Identification of intrinsic order and disorder in the DNA repair protein XPA. *Protein Sci* **10**, 560-571, doi:10.1110/ps.29401 (2001).

68 Buchko, G. W., Ni, S., Thrall, B. D. & Kennedy, M. A. Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA. *Nucleic Acids Res* **26**, 2779-2788 (1998).

69 Fadda, E. Role of the XPA protein in the NER pathway: A perspective on the function of structural disorder in macromolecular assembly. *Comput Struct Biotechnol J* **14**, 78-85, doi:10.1016/j.csbj.2015.11.007 (2016).

70 Vuzman, D. & Levy, Y. Intrinsically disordered regions as affinity tuners in protein-DNA interactions. *Mol Biosyst* **8**, 47-57, doi:10.1039/c1mb05273j (2012).

71 Dunker, A. K. *et al.* Protein disorder and the evolution of molecular recognition: theory, predictions and observations. *Pac Symp Biocomput*, 473-484 (1998).

72 Dunker, A. K. *et al.* Intrinsically disordered protein. *J Mol Graph Model* **19**, 26-59 (2001).

73 Tompa, P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett* **579**, 3346-3354, doi:10.1016/j.febslet.2005.03.072 (2005).

74 Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* **12**, 54-60 (2002).

75 Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* **33**, 2-8, doi:10.1016/j.tibs.2007.10.003 (2008).

76 Kuraoka, I. *et al.* Identification of a damaged-DNA binding domain of the XPA protein. *Mutat Res* **362**, 87-95 (1996).

77 Sugitani, N., Shell, S. M., Soss, S. E. & Chazin, W. J. Redefining the DNA-binding domain of human XPA. *J Am Chem Soc* **136**, 10830-10833, doi:10.1021/ja503020f (2014).

78 Tanaka, K. *et al.* Analysis of a human DNA excision repair gene involved in group A xeroderma pigmentosum and containing a zinc-finger domain. *Nature* **348**, 73-76, doi:10.1038/348073a0 (1990).

79 Asahina, H. *et al.* The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage. *Mutat Res* **315**, 229-237 (1994).

80 Hess, N. J. *et al.* Human nucleotide excision repair protein XPA: extended X-ray absorption fine-structure evidence for a metal-binding domain. *Protein Sci* **7**, 1970-1975, doi:10.1002/pro.5560070912 (1998).

81 Hu, J. *et al.* Metal binding mediated conformational change of XPA protein:a potential cytotoxic mechanism of nickel in the nucleotide excision repair. *J Mol Model* **22**, 156, doi:10.1007/s00894-016-3017-x (2016).

82 Buchko, G. W., Ni, S., Thrall, B. D. & Kennedy, M. A. Human nucleotide excision repair protein XPA: expression and NMR backbone assignments of the 14.7 kDa minimal damaged DNA binding domain (Met98-Phe219). *J Biomol NMR* **10**, 313-314 (1997).

83 Ikegami, T. *et al.* Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. *Nat Struct Biol* **5**, 701-706, doi:10.1038/1400 (1998).

84 Lian, F. M., Yang, X., Yang, W., Jiang, Y. L. & Qian, C. Structural characterization of the redefined DNA-binding domain of human XPA. *Biochem Biophys Res Commun*, doi:10.1016/j.bbrc.2019.05.050 (2019).

85 Koch, S. C. *et al.* Structural insights into the recognition of cisplatin and AAF-dG lesion by Rad14 (XPA). *Proceedings of the National Academy of Sciences of the United States of America* **112**, 8272-8277, doi:10.1073/pnas.1508509112 (2015).

86 Ebert, C., Simon, N., Schneider, S. & Carell, T. Structural Insights into the Recognition of N(2) -Aryl- and C8-Aryl DNA Lesions by the Repair Protein XPA/Rad14. *Chembiochem* **18**, 1379-1382, doi:10.1002/cbic.201700169 (2017).

87 Kokic, G. *et al.* Structural basis of TFIIH activation for nucleotide excision repair. *Nat Commun* **10**, 2885, doi:10.1038/s41467-019-10745-5 (2019).

88 Ikegami, T. *et al.* Resonance assignments, solution structure, and backbone dynamics of the DNA- and RPA-binding domain of human repair factor XPA. *J Biochem* **125**, 495-506 (1999).

89 Yang, Z. G., Liu, Y., Mao, L. Y., Zhang, J. T. & Zou, Y. Dimerization of human XPA and formation of XPA2-RPA protein complex. *Biochemistry* **41**, 13012-13020 (2002).

90 Liu, Y. *et al.* Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage. *Biochemistry* **44**, 7361-7368, doi:10.1021/bi047598y (2005).

91 Mustra, D. J., Warren, A. J., Wilcox, D. E. & Hamilton, J. W. Preferential binding of human XPA to the mitomycin C-DNA interstrand crosslink and modulation by arsenic and cadmium. *Chem Biol Interact* **168**, 159-168, doi:10.1016/j.cbi.2007.04.004 (2007).

92 Pradhan, S. & Mattaparthi, V. S. K. Structural dynamics and interactions of Xeroderma pigmentosum complementation group A (XPA98-210) with damaged DNA. *J Biomol Struct Dyn* **36**, 3341-3353, doi:10.1080/07391102.2017.1388285 (2018).

93 Robins, P., Jones, C. J., Biggerstaff, M., Lindahl, T. & Wood, R. D. Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA. *EMBO J* **10**, 3913-3921 (1991).

94 Jones, C. J. & Wood, R. D. Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochemistry* **32**, 12096-12104 (1993).

95 Iakoucheva, L. M., Walker, R. K., van Houten, B. & Ackerman, E. J. Equilibrium and stop-flow kinetic studies of fluorescently labeled DNA substrates with DNA repair proteins XPA and replication protein A. *Biochemistry* **41**, 131-143 (2002).

96 Hey, T., Lipps, G. & Krauss, G. Binding of XPA and RPA to damaged DNA investigated by fluorescence anisotropy. *Biochemistry* **40**, 2901-2910 (2001).

97 Wang, M., Mahrenholz, A. & Lee, S. H. RPA stabilizes the XPA-damaged DNA complex through protein-protein interaction. *Biochemistry* **39**, 6433-6439 (2000).

98 Buschta-Hedayat, N., Buterin, T., Hess, M. T., Missura, M. & Naegeli, H. Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6090-6095 (1999).

99 Brown, K. L. *et al.* Binding of the human nucleotide excision repair proteins XPA and XPC/HR23B to the 5R-thymine glycol lesion and structure of the cis-(5R,6S) thymine glycol epimer in the 5'-GTgG-3' sequence: destabilization of two base pairs at the lesion site. *Nucleic Acids Res* **38**, 428-440, doi:10.1093/nar/gkp844 (2010).

100 Yang, Z. *et al.* Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3'- and/or 5'-ssDNA branches. *Biochemistry* **45**, 15921-15930, doi:10.1021/bi061626q (2006).

101 Brabec, V., Stehlikova, K., Malina, J., Vojtiiskova, M. & Kasparkova, J. Thermodynamic properties of damaged DNA and its recognition by xeroderma pigmentosum group A protein and replication protein A. *Arch Biochem Biophys* **446**, 1-10, doi:10.1016/j.abb.2005.12.003 (2006).

102 Missura, M. *et al.* Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair. *EMBO J* **20**, 3554-3564, doi:10.1093/emboj/20.13.3554 (2001).

103 Camenisch, U., Dip, R., Schumacher, S. B., Schuler, B. & Naegeli, H. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nat Struct Mol Biol* **13**, 278-284, doi:10.1038/nsmb1061 (2006).

104 Stehlikova, K., Kostrhunova, H., Kasparkova, J. & Brabec, V. DNA bending and unwinding due to the major 1,2-GG intrastrand cross-link formed by antitumor cis-diamminedichloroplatinum(II) are flanking-base independent. *Nucleic Acids Res* **30**, 2894-2898, doi:10.1093/nar/gkf405 (2002).

105 Takahara, P. M., Rosenzweig, A. C., Frederick, C. A. & Lippard, S. J. Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. *Nature* **377**, 649-652, doi:10.1038/377649a0 (1995).

106 Kasparkova, J., Mellish, K. J., Qu, Y., Brabec, V. & Farrell, N. Site-specific d(GpG) intrastrand cross-links formed by dinuclear platinum complexes. Bending and NMR studies. *Biochemistry* **35**, 16705-16713, doi:10.1021/bi961160j (1996).

107 Kowalczyk, A. *et al.* Intrastrand DNA cross-links as tools for studying DNA replication and repair: two-, three-, and four-carbon tethers between the N(2) positions of adjacent guanines. *Biochemistry* **41**, 3109-3118, doi:10.1021/bi010450j (2002).

108 Shi, Y. B., Griffith, J., Gamper, H. & Hearst, J. E. Evidence for structural deformation of the DNA helix by a psoralen diadduct but not by a monoadduct. *Nucleic Acids Res* **16**, 8945-8952, doi:10.1093/nar/16.18.8945 (1988).

109 Spielmann, H. P., Dwyer, T. J., Hearst, J. E. & Wemmer, D. E. Solution structures of psoralen monoadducted and cross-linked DNA oligomers by NMR spectroscopy and restrained molecular dynamics. *Biochemistry* **34**, 12937-12953 (1995).

110 Mustra, D. J., Warren, A. J. & Hamilton, J. W. Preferential binding of human full-length XPA and the minimal DNA binding domain (XPA-MF122) with the mitomycin C-DNA interstrand cross-link. *Biochemistry* **40**, 7158-7164 (2001).

111 Warren, A. J. & Hamilton, J. W. Synthesis and structural characterization of the N2G-mitomycin C-N2G interstrand cross-link in a model synthetic 23 base pair oligonucleotide DNA duplex. *Chem Res Toxicol* **9**, 1063-1071, doi:10.1021/tx960070c (1996).

112 Sugitani, N., Voehler, M. W., Roh, M. S., Topolska-Wos, A. M. & Chazin, W. J. Analysis of DNA binding by human factor xeroderma pigmentosum complementation group A (XPA) provides insight into its interactions with nucleotide excision repair substrates. *J Biol Chem* **292**, 16847-16857, doi:10.1074/jbc.M117.800078 (2017).

113 Wakasugi, M. *et al.* Physical and functional interaction between DDB and XPA in nucleotide excision repair. *Nucleic Acids Res* **37**, 516-525, doi:10.1093/nar/gkn964 (2009).

114 You, J. S., Wang, M. & Lee, S. H. Biochemical analysis of the damage recognition process in nucleotide excision repair. *J Biol Chem* **278**, 7476-7485, doi:10.1074/jbc.M210603200 (2003).

115 Bunick, C. G., Miller, M. R., Fuller, B. E., Fanning, E. & Chazin, W. J. Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein. *Biochemistry* **45**, 14965-14979, doi:10.1021/bi061370o (2006).

116 Krasikova, Y. S. *et al.* Interaction of nucleotide excision repair factors XPC-HR23B, XPA, and RPA with damaged DNA. *Biochemistry (Mosc)* **73**, 886-896 (2008).

117 Pradhan, S., Das, P. & Mattaparthi, V. S. K. Characterizing the Binding Interactions between DNA-Binding Proteins, XPA and XPE: A Molecular Dynamics Approach. *ACS Omega* **3**, 15442-15454, doi:10.1021/acsomega.8b01793 (2018).

118 Park, C. H., Mu, D., Reardon, J. T. & Sancar, A. The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIE transcription factor. *J Biol Chem* **270**, 4896-4902 (1995).

119 Ziani, S. *et al.* Sequential and ordered assembly of a large DNA repair complex on undamaged chromatin. *J Cell Biol* **206**, 589-598, doi:10.1083/jcb.201403096 (2014).

120 He, Z., Henricksen, L. A., Wold, M. S. & Ingles, C. J. RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature* **374**, 566-569, doi:10.1038/374566a0 (1995).

121 Lee, B. E. *et al.* Functional studies on the interaction between human replication protein A and Xeroderma pigmentosum group A complementing protein (XPA). *Mol Cells* **9**, 185-190 (1999).

122 Saijo, M., Kuraoka, I., Masutani, C., Hanaoka, F. & Tanaka, K. Sequential binding of DNA repair proteins RPA and ERCC1 to XPA in vitro. *Nucleic Acids Res* **24**, 4719-4724 (1996).

123 Lee, S. H., Kim, D. K. & Drissi, R. Human xeroderma pigmentosum group A protein interacts with human replication protein A and inhibits DNA replication. *J Biol Chem* **270**, 21800-21805, doi:10.1074/jbc.270.37.21800 (1995).

124 Daughdrill, G. W. *et al.* Chemical shift changes provide evidence for overlapping single-stranded DNA- and XPA-binding sites on the 70 kDa subunit of human replication protein A. *Nucleic Acids Res* **31**, 4176-4183 (2003).

125 McNeil, E. M. & Melton, D. W. DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. *Nucleic Acids Res* **40**, 9990-10004, doi:10.1093/nar/gks818 (2012).

126 Nagai, A. *et al.* Enhancement of damage-specific DNA binding of XPA by interaction with the ERCC1 DNA repair protein. *Biochem Biophys Res Commun* **211**, 960-966 (1995).

127 Pei, J., Kim, B. H. & Grishin, N. V. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* **36**, 2295-2300, doi:10.1093/nar/gkn072 (2008).

128 Romero, P. *et al.* Sequence complexity of disordered protein. *Proteins* **42**, 38-48 (2001).

129 Beckwitt, E. C., Kong, M. & Van Houten, B. Studying protein-DNA interactions using atomic force microscopy. *Semin Cell Dev Biol* **73**, 220-230, doi:10.1016/j.semcdb.2017.06.028 (2018).

130 Binnig, G., Quate, C. F. & Gerber, C. Atomic Force Microscope. *Physical Review Letters* **56**, 930-933 (1986).

131 Sauer, R. T. in *Methods in Enzymology* Vol. 208 (Academic Press, San Diego, California, 1991).

132 Jen-Jacobson, L., Engler, L. E. & Jacobson, L. A. Structural and thermodynamic strategies for site-specific DNA binding proteins. *Structure* **8**, 1015-1023 (2000).

133 Bustamante, C. *et al.* Circular DNA molecules imaged in air by scanning force microscopy. *Biochemistry* **31**, 22-26 (1992).

134 Yang, J., Takeyasu, K. & Shao, Z. Atomic force microscopy of DNA molecules. *FEBS Lett* **301**, 173-176 (1992).

135 Klein, H. L. *et al.* Guidelines for DNA recombination and repair studies: Mechanistic assays of DNA repair processes. *Microb Cell* **6**, 65-101, doi:10.15698/mic2019.01.665 (2019).

136 Kad, N. M. & Van Houten, B. in *DNA Repair (Amst)* Vol. 20 1-153 (2014).

137 Monachino, E., Spenkelink, L. M. & van Oijen, A. M. Watching cellular machinery in action, one molecule at a time. *J Cell Biol* **216**, 41-51, doi:10.1083/jcb.201610025 (2017).

138 Kad, N. M., Wang, H., Kennedy, G. G., Warshaw, D. M. & Van Houten, B. Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by single-molecule imaging of quantum-dot-labeled proteins. *Mol Cell* **37**, 702-713, doi:10.1016/j.molcel.2010.02.003 (2010).

139 Van Houten, B. & Kad, N. Investigation of bacterial nucleotide excision repair using single-molecule techniques. *DNA Repair (Amst)* **20**, 41-48, doi:10.1016/j.dnarep.2013.10.012 (2014).

140 Dunn, A. R., Kad, N. M., Nelson, S. R., Warshaw, D. M. & Wallace, S. S. Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. *Nucleic Acids Res* **39**, 7487-7498, doi:10.1093/nar/gkr459 (2011).

141 Graneli, A., Yeykal, C. C., Robertson, R. B. & Greene, E. C. Long-distance lateral diffusion of human Rad51 on double-stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 1221-1226, doi:10.1073/pnas.0508366103 (2006).

142 Kong, M., Beckwitt, E. C., Springall, L., Kad, N. M. & Van Houten, B. Single-Molecule Methods for Nucleotide Excision Repair: Building a System to Watch Repair in Real Time. *Methods Enzymol* **592**, 213-257, doi:10.1016/bs.mie.2017.03.027 (2017).

143 Schaefer, C. M. *et al.* FadA5 a thiolase from Mycobacterium tuberculosis: a steroid-binding pocket reveals the potential for drug development against tuberculosis. *Structure* **23**, 21-33, doi:10.1016/j.str.2014.10.010 (2015).

144 Sauer, F. *et al.* Differential Oligomerization of the Deubiquitinases USP25 and USP28 Regulates Their Activities. *Mol Cell* **74**, 421-435 e410, doi:10.1016/j.molcel.2019.02.029 (2019).

145 Schofield, M. J., Lilley, D. M. & White, M. F. Dissection of the sequence specificity of the Holliday junction endonuclease CCE1. *Biochemistry* **37**, 7733-7740, doi:10.1021/bi980399s (1998).

146 Pollard, T. D. A guide to simple and informative binding assays. *Mol Biol Cell* **21**, 4061-4067, doi:10.1091/mbc.E10-08-0683 (2010).

147 Luo, J. *et al.* Optical Control of DNA Helicase Function through Genetic Code Expansion. *Chembiochem* **18**, 466-469, doi:10.1002/cbic.201600624 (2017).

148 Wang, H., Nora, G. J., Ghodke, H. & Opresko, P. L. Single molecule studies of physiologically relevant telomeric tails reveal POT1 mechanism for promoting G-quadruplex unfolding. *J Biol Chem* **286**, 7479-7489, doi:10.1074/jbc.M110.205641 (2011).

149 Yang, Y., Sass, L. E., Du, C., Hsieh, P. & Erie, D. A. Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions. *Nucleic Acids Res* **33**, 4322-4334, doi:10.1093/nar/gki708 (2005).

150 Baldwin, M. R. & O'Brien, P. J. Nonspecific DNA binding and coordination of the first two steps of base excision repair. *Biochemistry* **49**, 7879-7891, doi:10.1021/bi100889r (2010).

151 Nelson, S. R., Dunn, A. R., Kathe, S. D., Warshaw, D. M. & Wallace, S. S. Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2091-2099, doi:10.1073/pnas.1400386111 (2014).

152 Arnspang, E. C., Brewer, J. R. & Lagerholm, B. C. Multi-color single particle tracking with quantum dots. *PLoS One* **7**, e48521, doi:10.1371/journal.pone.0048521 (2012).

153 Schurr, J. M. The one-dimensional diffusion coefficient of proteins absorbed on DNA. Hydrodynamic considerations. *Biophys Chem* **9**, 413-414 (1979).

154 Bagchi, B., Blainey, P. C. & Xie, X. S. Diffusion constant of a nonspecifically bound protein undergoing curvilinear motion along DNA. *J Phys Chem B* **112**, 6282-6284, doi:10.1021/jp077568f (2008).

155 Kaemmer, S. B. Introduction to bruker's scanasyst and peakforce tapping afm technology. *Bruker application note. Bruker Nano Inc., Santa Barbara, CA* (2011).

156 Pyne, A., Thompson, R., Leung, C., Roy, D. & Hoogenboom, B. W. Single-molecule reconstruction of oligonucleotide secondary structure by atomic force microscopy. *Small* **10**, 3257-3261, doi:10.1002/smll.201400265 (2014).

157 Murugesapillai, D. *et al.* DNA bridging and looping by HMO1 provides a mechanism for stabilizing nucleosome-free chromatin. *Nucleic Acids Res* **42**, 8996-9004, doi:10.1093/nar/gku635 (2014).

158 Eker, A. P. *et al.* Xeroderma pigmentosum group A correcting protein from calf thymus. *Mutat Res* **274**, 211-224 (1992).

159 Datta, H. J., Chan, P. P., Vasquez, K. M., Gupta, R. C. & Glazer, P. M. Triplex-induced recombination in human cell-free extracts. Dependence on XPA and HsRad51. *J Biol Chem* **276**, 18018-18023, doi:10.1074/jbc.M011646200 (2001).

160 Schneider, S. W., Larmer, J., Henderson, R. M. & Oberleithner, H. Molecular weights of individual proteins correlate with molecular volumes measured by atomic force microscopy. *Pflugers Arch* **435**, 362-367, doi:10.1007/s004240050524 (1998).

161 Ratcliff, G. C. & Erie, D. A. A Novel Single-Molecule Study To Determine Protein−Protein Association Constants. *Journal of the American Chemical Society* **123**, 5632-5635, doi:10.1021/ja005750n (2001).

162 Rademakers, S. *et al.* Xeroderma pigmentosum group A protein loads as a separate factor onto DNA lesions. *Mol Cell Biol* **23**, 5755-5767 (2003).

163 Pratto, F., Suzuki, Y., Takeyasu, K. & Alonso, J. C. Single-molecule analysis of protein-DNA complexes formed during partition of newly replicated plasmid molecules in Streptococcus pyogenes. *J Biol Chem* **284**, 30298-30306, doi:10.1074/jbc.M109.035410 (2009).

164 Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J. & Pelletier, H. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**, 11205-11215, doi:10.1021/bi9703812 (1997).

165 O'Handley, S. F. *et al.* Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics. *Biochemistry* **32**, 2481-2497 (1993).

166 Grad, R., Shapiro, R., Hingerty, B. E. & Broyde, S. A molecular mechanics and dynamics study of the minor adduct between DNA and the carcinogen 2-(acetylamino)fluorene (dG-N2-AAF). *Chem Res Toxicol* **10**, 1123-1132, doi:10.1021/tx970092e (1997).

167 Kong, M. & Van Houten, B. Rad4 recognition-at-a-distance: Physical basis of conformation-specific anomalous diffusion of DNA repair proteins. *Prog Biophys Mol Biol* **127**, 93-104, doi:10.1016/j.pbiomolbio.2016.12.004 (2017).

168 Lin, J. *et al.* Functional interplay between SA1 and TRF1 in telomeric DNA binding and DNA-DNA pairing. *Nucleic Acids Res* **44**, 6363-6376, doi:10.1093/nar/gkw518 (2016).

169 Tan, C., Terakawa, T. & Takada, S. Dynamic Coupling among Protein Binding, Sliding, and DNA Bending Revealed by Molecular Dynamics. *J Am Chem Soc* **138**, 8512-8522, doi:10.1021/jacs.6b03729 (2016).

170 Yeh, J. I. *et al.* Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2737-2746, doi:10.1073/pnas.1110067109 (2012).

171 Collins, B. E., Ye, L. F., Duzdevich, D. & Greene, E. C. DNA curtains: novel tools for imaging protein-nucleic acid interactions at the single-molecule level. *Methods Cell Biol* **123**, 217-234, doi:10.1016/B978-0-12-420138-5.00012-4 (2014).

172 Fazio, T., Visnapuu, M. L., Wind, S. & Greene, E. C. DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging. *Langmuir* **24**, 10524-10531, doi:10.1021/la801762h (2008).

173 Visnapuu, M. L., Fazio, T., Wind, S. & Greene, E. C. Parallel arrays of geometric nanowells for assembling curtains of DNA with controlled lateral dispersion. *Langmuir* **24**, 11293-11299, doi:10.1021/la8017634 (2008).

174 Brown, M. W. *et al.* Dynamic DNA binding licenses a repair factor to bypass roadblocks in search of DNA lesions. *Nat Commun* **7**, 10607, doi:10.1038/ncomms10607 (2016).

175 Li, L., Peterson, C. A., Lu, X. & Legerski, R. J. Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Mol Cell Biol* **15**, 1993-1998 (1995).

176 Tripsianes, K. *et al.* Analysis of the XPA and ssDNA-binding surfaces on the central domain of human ERCC1 reveals evidence for subfunctionalization. *Nucleic Acids Res* **35**, 5789-5798, doi:10.1093/nar/gkm503 (2007).

177 Park, C. H. & Sancar, A. Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5017-5021 (1994).

178 Tsodikov, O. V. *et al.* Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA. *EMBO J* **26**, 4768-4776, doi:10.1038/sj.emboj.7601894 (2007).

179 Wu, X., Shell, S. M., Liu, Y. & Zou, Y. ATR-dependent checkpoint modulates XPA nuclear import in response to UV irradiation. *Oncogene* **26**, 757-764, doi:10.1038/sj.onc.1209828 (2007).

180 Shell, S. M. *et al.* Checkpoint kinase ATR promotes nucleotide excision repair of UV-induced DNA damage via physical interaction with xeroderma pigmentosum group A. *J Biol Chem* **284**, 24213-24222, doi:10.1074/jbc.M109.000745 (2009).

181 Lange, S. S., Reddy, M. C. & Vasquez, K. M. Human HMGB1 directly facilitates interactions between nucleotide excision repair proteins on triplex-directed psoralen interstrand crosslinks. *DNA Repair (Amst)* **8**, 865-872, doi:10.1016/j.dnarep.2009.04.001 (2009).

182 Mukherjee, A. & Vasquez, K. M. HMGB1 interacts with XPA to facilitate the processing of DNA interstrand crosslinks in human cells. *Nucleic Acids Res* **44**, 1151-1160, doi:10.1093/nar/gkv1183 (2016).

183 King, B. S., Cooper, K. L., Liu, K. J. & Hudson, L. G. Poly(ADP-ribose) contributes to an association between poly(ADP-ribose) polymerase-1 and xeroderma pigmentosum complementation group A in nucleotide excision repair. *J Biol Chem* **287**, 39824-39833, doi:10.1074/jbc.M112.393504 (2012).

184 Nitta, M. *et al.* A novel cytoplasmic GTPase XAB1 interacts with DNA repair protein XPA. *Nucleic Acids Res* **28**, 4212-4218 (2000).

185 Nakatsu, Y. *et al.* XAB2, a novel tetratricopeptide repeat protein involved in transcription-coupled DNA repair and transcription. *J Biol Chem* **275**, 34931-34937, doi:10.1074/jbc.M004936200 (2000).