SINGLE-MOLECULE STUDIES OF RAD4-RAD23 REVEAL A DYNAMIC DNA DAMAGE RECOGNITION PROCESS

by

Muwen Kong

B.A. Physics, Knox College, 2008

M.S. Physics, Carnegie Mellon University, 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

2017

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Muwen Kong

It was defended on

June 30, 2017

and approved by

Guillermo Romero, PhD., Associate Professor, Department of Pharmacology and Chemical Biology

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

Neil Kad, PhD., Senior Lecturer, School of Biosciences, University of Kent

Patricia Opresko, PhD., Associate Professor, Department of Environmental and Occupational Health

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

Single-Molecule Studies of Rad4-Rad23 Reveal a Dynamic DNA Damage Recognition Process

Muwen Kong, PhD

University of Pittsburgh, 2017

Nucleotide excision repair (NER) is an evolutionarily conserved mechanism that processes helix-destabilizing and/or -distorting DNA lesions, such as UV-induced photoproducts. As the first step towards productive repair, the human NER damage sensor XPC-RAD23B needs to efficiently locate sites of damage among billons of base pairs of undamaged DNA. In this dissertation, we investigated the dynamic protein-DNA interactions during the damage recognition step using a combination of fluorescence-based single-molecule DNA tightrope assays, atomic force microscopy, as well as cell survival and *in vivo* repair kinetics assays. We observed that quantum dot-labeled Rad4-Rad23, the yeast homolog of human XPC-RAD23B, formed nonmotile complexes on DNA or conducted a one-dimensional search via either random diffusion or constrained motion along DNA. Using atomic force microscopy, we studied binding of Rad4 lacking the β-hairpin domain 3 (BHD3) to damage-containing DNA and found that this structural motif is non-essential for damage-specific binding or DNA bending. Furthermore, we demonstrated that deletion of seven residues in the tip of β-hairpin in BHD3 increased Rad4-Rad23 constrained motion at the expense of stable binding at sites of DNA lesions, without diminishing cellular UV resistance or photoproduct repair *in vivo*. These results suggest a distinct intermediate in the damage recognition process during NER, allowing dynamic DNA damage detection at a distance. Finally, we explore existing physical models and examples of subdiffusive motion, and discuss a model in which constrained motion by Rad4-Rad23 on DNA may be driven by conformational changes of the protein.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

What I cannot create, I do not understand.

-Richard Feynman

I will be the first one to admit that I never thought that I would get here, writing about the body of work that I can proudly call my own. As Matt Might, professor of Computer Science at University of Utah, so eloquently illustrated in *The Illustrated Guide to Ph.D.*, you keep pushing the boundary of human knowledge for a few years, before eventually, the boundary gives way and you make the tiniest dent in it, and that dent is called a Ph. D. Through my journey to reach and eventually bang my head against said boundary, I have been fortunate to interact with great mentors and colleagues, without whom this almost certainly would have been mission impossible.

I am grateful for my earlier education and research experience at Knox College Department of Physics, particularly under the mentorship of Chuck Schulz, who allowed me to first dip into the field of biophysical research. I also appreciate the opportunity to have spent time in the laboratory of Maumita Mandal at Carnegie Mellon University. Even though regrettably the road was not without bumps and turns, I certainly benefited greatly from having Guangtao Song and Huizhong Xu as friends and colleagues.

I have been very fortunate indeed to finally settle in the laboratory of Bennett Van Houten. I could not have asked for a better mentor for guidance during my time in graduate school, and hopefully for years to come. Ben encourages the students to think on their own and treats them as his colleagues. His breadth of knowledge and willingness to learn from others serve as daily

inspirations for me. Ben's passion for science, devotion to thought experiments, and mantra of 'if you don't have time to do it right, when do you have time to do it again' propel me and the rest of the group forward. Thank you for being the most enthusiastic cheerleader and the harshest critic at the same time for the past few years.

I am also thankful for having friends and colleagues who made times spent in and out of the lab more productive and enjoyable: Harshad Ghodke, Lili Liu, Emily Beckwitt, Sunbok Jang, Vera Roginskaya, and the rest of the Van Houten laboratory; our undergraduate and rotation students including Katee Driscoll, Ananya Mukundan, Caitlin Johnson, Izzy Carnaval; members of the Opresko, Bakkenist, Lan, Bernstein, and Sobol laboratories; and friends from the MBSB program, past and present, including Abhishek Mandal, Sean Carney, and Ryan Slack.

In addition, I would like to thank my thesis Committee, Guillermo Romero, Marcel Bruchez, Neil Kad, and Patty Opresko, for guiding the project and asking the tough questions along the way. I also thank Simon Watkins and Greg Gibson at CBI, Jung-Hyun Min and Xuejing Chen at the University of Illinois at Chicago, and John Wyrick and Peng Mao at the Washington State University, for a fruitful collaboration and lending their expertise where I fall short.

Finally, I certainly would not have been where I am without the support of my parents, who have stood by me through all the highs and lows since I decided, almost 15 years ago, to pursue my higher education halfway across the globe. And lastly, I want to thank Beijing for always being there and having the faith and patience in me.

1.0 INTRODUCTION

Despite the continuous assaults suffered from endogenous and exogenous agents, all living organisms on this planet, regardless of how brief their lifespan, strive to pass on their complete and intact genetic information to the next generation. With genome integrity constantly under threat, a complex network of DNA repair pathways, as a part of the sophisticated DNA damage response, has emerged and evolved over billions of years. While each mammalian cell encounters ~10^4 – 10^5 DNA lesions per day (1), the system of distinct yet complementary repair mechanisms allows efficient recognition and repair of an incredibly wide spectrum of DNA damage. These lesions range from oxidative or alkylating base damage and mismatched bases, to bulky DNA adducts and single- or double-stranded breaks. The Nobel Prize in Chemistry 2015 was awarded to Thomas Lindahl, Paul Modrich, and Aziz Sancar, in recognition of their pioneering work on dissecting the molecular mechanisms of base excision repair (BER), mismatch repair (MMR), and nucleotide excision repair (NER), respectively.

1.1 NUCLEOTIDE EXCISION REPAIR

The year 1964 marked a milestone for the emerging field of nucleotide excision repair (NER). Richard Setlow, Paul Swenson, and William Carrier (Oak Ridge National Laboratory), Paul Howard-Flanders and his postdoctoral fellow Richard Boyce (Yale University), as well as Philip

Hanawalt, a former graduate student of Setlow's, and his graduate student David Pettijohn (Stanford University) independently reported findings that elucidated aspects of a novel DNA repair mechanism for UV-induced DNA lesions in *E. coli* (2). In the decade following the discovery of enzymatic photoreactivation in 1949 by Renato Dulbecco (3) and Albert Kelner (4), the notion of a putative light-independent dark repair was suggested by Robert Haynes and others, based on the observation that survival and yield of mutations in UV-irradiated cells were influenced by post-treatment conditions (5). Subsequently in 1964, seminal work by Setlow and Carrier showed that UV-induced thymine dimers were removed in a time-dependent manner from acid-insoluble high molecular weight DNA in the form of short oligonucleotides that were recovered from the acid-soluble fraction from bacteria (6). This finding was corroborated by a similar observation made independently by Howard-Flanders and Boyce using the *K-12* strain of *E. coli* (7). At the same time, using tritium-labeled 5-bromouracil followed by CsCl density gradient to examine DNA replication in *E. coli* post-UV-irradiation, Hanawalt and Pettijohn demonstrated non-conservative synthesis of short single-stranded DNA (ssDNA) (8). Even though the molecular underpinnings of bacterial nucleotide excision repair would not be delineated for another two decades through the work of Aziz Sancar and others, the discoveries made by the pioneering laboratories of Setlow, Howard-Flanders, and Hanawalt shed the first rays of light on the mysteries of dark repair.

1.1.1 Overview

The earliest evidence of NER in mammalian cells was presented by Robert Painter and Ronald Rasmussen shortly after the discovery of bacterial NER (9). Soon after, James Cleaver showed that cells from patients with the hereditary disease xeroderma pigmentosum (XP) (see 1.2.1 below)

were deficient in repair synthesis after UV irradiation, strongly suggesting a connection between the genetic disorder and deficiencies in NER (10). In the five decades that followed, generations of researchers dedicated their careers to dissecting the mechanism of NER and understanding its roles in cancer.

Ultraviolet (UV) light from solar radiation is the most pervasive environmental DNA-damaging agent, capable of inducing ~100,000 lesions per cell per hour (11). UV irradiation of DNA produces predominantly two types of lesions, cyclobutane pyrimidine dimers (CPDs) and 6-4 pyrimidine-pyrimidone photoproducts (6-4PPs), at a ratio of approximately 3 CPD : 1 6-4PP (12). While NER is primarily tasked with the repair of UV-induced photoproducts, the pathway is capable of recognizing a broad spectrum of bulky DNA lesions including liver carcinogen 2-acetylaminofluorene (AAF) adducts (13), psoralen mono-adducts (14), anti-cancer drug cisplatin intra-strand crosslinks (15), environmental mutagen polycyclic aromatic hydrocarbon benzo[α]pyrene diol epoxide (BPDE) adducts (16), among many others (17).

NER can be initiated via two separate mechanisms, transcription-coupled NER (TC-NER) (Figure 1.1) or global genome NER (GG-NER) (Figure 1.2), depending on how DNA damage is recognized. These two sub-pathways then converge and proceed through damage verification, assembly of incision complex, excision, repair synthesis and ligation (Figure 1.3). Although the general principle of 'recognize, remove, and restore' applies to both prokaryotic and eukaryotic NER, the latter is significantly more complicated and tightly regulated: the eukaryotic GG-NER is orchestrated by over 30 different factors whereas the same process in prokaryotic cells can be carried out by only 6 proteins. The following sections provide a review of the current understanding and model of the eukaryotic NER pathway.

1.1.2 Initiation of TC-NER

First observed in the laboratory of Philip Hanawalt (18), TC-NER serves to preferentially initiate repair of RNA polymerase-blocking lesions in the transcribed strand, allowing interrupted transcription and gene expression to resume (Figure 1.1). In mammalian cells, damage recognition in TC-NER is effectively accomplished by the way of a lesion-stalled RNA polymerase II (RNA Pol II) (19). Cockayne syndrome protein B (CSB, ERCC6) is a DNA-dependent ATPase that translocates along DNA with the transcription machinery (20). Upon transcription arrest at the site of lesion, interaction of CSB with the transcription elongation complex is stabilized (21), while CSB actively changes DNA conformation by wrapping DNA around itself (22). At the site of stalled RNA Pol II, CSB also functions to recruit other core NER factors, the Cockayne syndrome WD repeat protein A (CSA, ERCC8) complex, and histone acetyltransferase p300 (23). The CSA complex, an E3 ubiquitin ligase ($CRL4^{CSA}$) consisting of CSA, DDB1, cullin 4A (Cul4A), and regulator of cullins 1 (Roc1/Rbx1), associates with COP9 signalosome (CSN) (24) and regulates degradation of its substrate CSB through the ubiquitin-proteasome pathway (25). Nonetheless, in cooperation with CSB, the CSA complex is indispensable for recruitment of transcription factor IIS (TFIIS), XPA Binding Protein 2 (XAB2), and high mobility group nucleosome-binding domain-containing protein 1 (HMGN1) (23). More recently, it's been shown that CSA also recruits the UV-stimulated scaffold protein A (UVSSA) following UV-irradiation (26). UVSSA interacts with the elongating RNA Pol II, and together with its interaction partner ubiquitin carboxyl-terminal hydrolase 7 (USP7), work to promote stabilization of CSB via deubiquitination (27-29). With the major protein components identified, it remains unclear how the arrested RNA Pol II, which occupies a footprint of ~35 nucleotides around the lesion (30), allows access of the lesion by the core NER complex (31). To date, several mechanisms have been proposed, including: 1)

TFIIS-mediated backtracking of RNA Pol II, which may require upstream nucleosome sliding facilitated by HMGN1 and p300 (30); 2) transcription factor IIH (TFIIH)-induced phosphorylation-dependent conformational change of RNA Pol II, through the cooperative recognition of the stalled polymerase and transcription bubble by XPG and CSB (32); and 3) poly-ubiquitination and proteasome degradation of RNA Pol II, perhaps as a last resort (33).

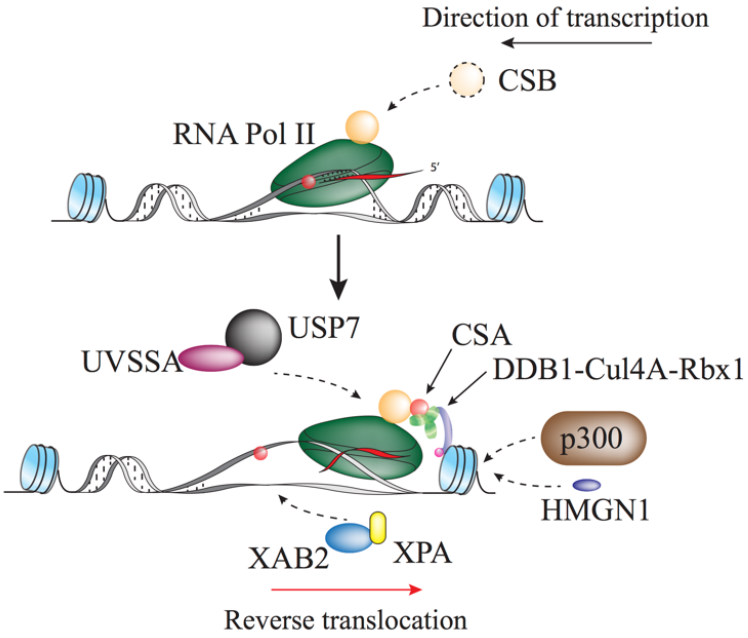

Figure 1.1: Initiation of transcription-coupled repair.

Adapted with permission from (34).

1.1.3 Initiation of GG-NER

For the rest of the genome that is not being actively transcribed, initiation of GG-NER relies on damage recognition through specific binding of the UV-damaged DNA-binding protein (UV-DDB) complex or the heterotrimeric complex XPC-RAD23B-CETN2 (35) (Figure 1.2).

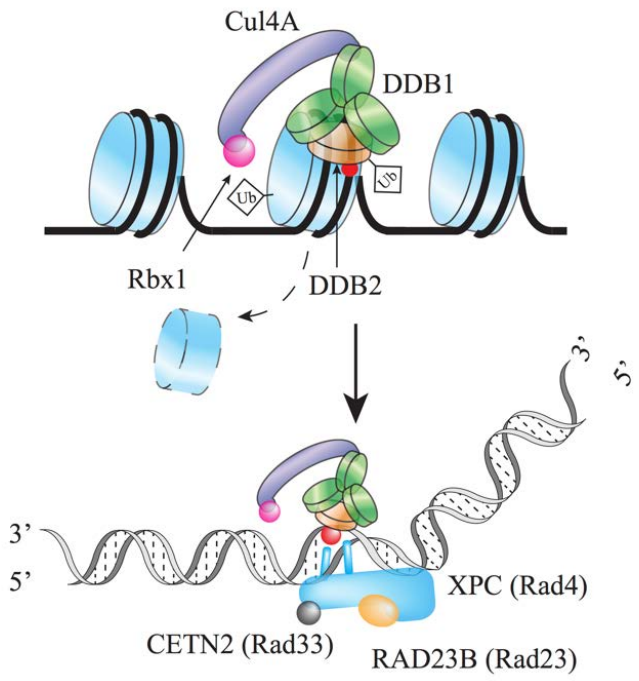

Figure 1.2: Initiation of global genome repair.

Adapted with permission from (34).

To ensure efficient repair of CPDs, the UV-DDB complex, first arrives at sites of damage and stimulates repair by recruiting XPC (36-40). UV-DDB, consisted of DDB1 and DDB2 subunits (41), binds tightly to CPDs and 6-4PPs as well as abasic sites and mismatches (42). It has been shown that upon damage recognition through multiple kinetic intermediates and conformational proofreading, UV-DDB forms stable dimer of dimers on DNA at sites of lesion (43,44). In addition, UV-DDB also forms an E3 ubiquitin ligase complex $CRL4^{DDB2}$ with Cul4A and Roc1/Rbx1 (45). Like the $CRL4^{CSA}$ complex, the E3 ligase activity of $CRL4^{DDB2}$ activated by the dissociation of CSN and neddylation of Cul4 by the ubiquitin-like protein NEDD8 (24,46). Substrates targeted for ubiquitination by the $CRL4^{DDB2}$ E3 ligase include the DDB2 subunit, XPC, as well as members of the histone octamer for nucleosome relaxation (47-51). UV-induced $CLR4^{DDB2}$-mediatated ubiquitination of XPC has been suggested to play a critical role in the

efficient lesion hand-off from UV-DDB to XPC (52), degradation of XPC for recruitment of downstream NER factors (53,54), and protection of DDB2 from excessive ubiquitination and degradation (55). Other post-translational modifications that regulate DDB2 stability include deubiquitination by USP24 (56) and poly-ADP-ribosylation (PARylation) by poly(ADP-ribose) polymerase 1 (PARP1) (57,58).

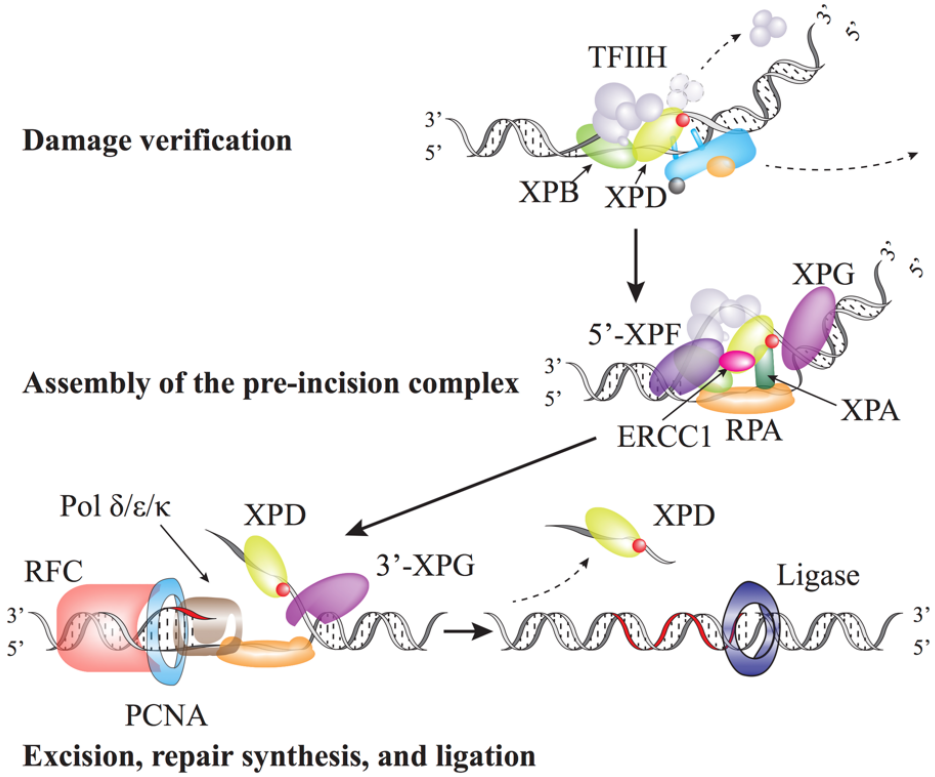

Figure 1.3: Post-damage recognition steps of mammalian NER.

Adapted with permission from (34).

In association with XPC (59,60), RAD23A/B are two human homologs of the *S. cerevisiae* Rad23 (60) and functionally interchangeable in NER. They both stimulate repair (61,62) and protect XPC from proteasome degradation (63,64). While both RAD23A and RAD23B associate with XPC with similar affinities, RAD23B is ~10-fold more abundant *in vivo*, therefore leading to mostly XPC-RAD23B complexes being observed (64). Besides their roles in NER, RAD23A/B

also participate in cellular functions such as protein degradation as well as cell cycle control and apoptosis (65). Centrin-2 (CETN2), a calcium-binding protein also involved in centrosome-related structures, binds to XPC through a domain near the disordered C-terminus of XPC and stimulates NER (66-70). The XPC complex has long been considered the main initiator of GG-NER (71) and its binding to damaged DNA, which coincides with the dissociation of RAD23B (72), a prerequisite for subsequent assembly of core NER factors (73-75). Consistent with the versatile role that NER plays in removing structurally distinct lesions, early studies showed XPC binds avidly to 6-4PP, BPDE, cholesterol, cisplatin adducts, thymine glycol, small DNA bubbles and loops, but curiously not CPDs (76-84). To explain the diversity in the chemical structures of binding substrates, it was suggested that, as a first step in damage recognition, XPC specifically binds to locally destabilized DNA helical regions (81,85). Such proposition was corroborated by the co-crystal structure of Rad4-Rad23 (Figure 1.4), the *S. cerevisiae* homologs of human XPC-RAD23B (60,86,87), in complex with a 24-bp DNA duplex harboring a CPD lesion opposite of a pair of mismatched thymidines (88). Structural evidence reveals that Rad4 inserts a β-hairpin between the two strands of DNA at the site of CPD, causing the opposite bases in the undamaged strand to flip out, which are then stabilized by the hydrophobic amino acids in the protein (88). Furthermore, the Rad4-DNA structure helps explain the lack of specificity of XPC towards CPD, for CPD alone only causes mild distortions in the DNA helix (89). A recent high-throughput affinity-purification mass spectrometry study has identified physical interactions between XPC-RAD23B and PARP1/2 (90). PARylation of both subunits of the complex by PARP1 has been demonstrated (91). Other previously identified post-translational modifications of XPC also include ubiquitination by RNF111 and small ubiquitin-like modifier (SUMO)-ylations (53,54,92-94).

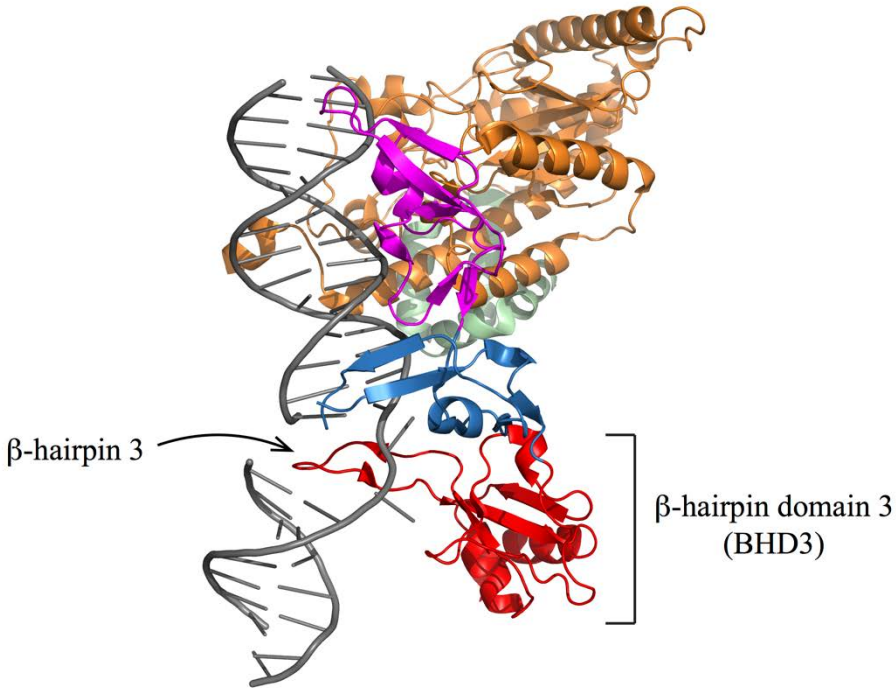

Figure 1.4: Crystal structure of Rad4-Rad23 in complex with CPD-containing DNA duplex (PDB: 2QSG).

1.1.4 Damage Verification and Assembly of Pre-Incision Complex

TFIIH can be recruited to the site of lesion through interactions with both the N- and C-terminal regions of XPC (73,74,95-97). TFIIH, a 10-subunit general transcription factor consisting of the 7-member core (XPB, XPD, p8/TTDA, p34, p44, p52, and p62) and the trimeric cyclin-dependent kinase (CDK)-activating kinase (CAK) complex (CyclinH, CDK7, MAT1), is involved in both basal transcription and NER (98). Both XPB and p62 subunits have been implicated in TFIIH interactions with and recruitment by XPC (95,99,100). CAK, while essential for transcription initiation, negatively regulates NER (101), therefore must dissociate from TFIIH during repair (102). XPB (ERCC3, 3'-5') and XPD (ERCC2, 5'-3') are helicases of opposite polarities (98). It has been suggested that the ATPase activity and a subsequent conformational change of XPB that

result in DNA opening are essential in recruitment and positioning of TFIIH (75,103,104). In contrast, the helicase activity of XPD is believed to be indispensable for NER (105) and responsible for the observed 5'-3' scanning by TFIIH on DNA (106). Indeed, crystal structures of XPD and its homologs revealed that the protein contains an ARCH domain and a 4Fe-S cluster, which form an opening sufficient only for ssDNA to pass through (107-110), supporting the hypothesis where damage verification involves stalling of a translocating XPD at the site of lesion (106,111). The ARCH domain plays a role in anchoring CAK to TFIIH (112). The presence of CAK negatively regulates the helicase activity of XPD and therefore inhibits NER (112), consistent with the previous finding that NER is driven by dissociation of CAK (102). With DNA partially unwound for damage verification by TFIIH, replication protein A (RPA) and XPA are recruited to protect exposed ssDNA and further stabilize the complex (74,113). Now considered a central player in coordinating the NER complex, XPA not only interacts with almost every NER factor (114), including DDB2 (115), XPC-RAD23B (116,117), TFIIH (118-120), RPA (121-125), XPF-ERCC1 (126-128), and PCNA (129), but also recognizes certain lesions and structural features in DNA (130-136). The importance of XPA during damage verification was highlighted by a recent study where XPA was found to effectively stimulate TFIIH scanning on undamaged DNA and promote TFIIH stalling in the presence of a bulky lesion (137). With XPA in place at the 5' side of the TFIIH-induced DNA bubble and RPA cooperatively bound to undamaged ssDNA (138), structure-specific nucleases XPF-ERCC1 and XPG can be oriented (139) and recruited to the complex by XPA (126-128) and TFIIH (97,140-142), respectively.

1.1.5 Excision, Repair Synthesis, and Ligation

Dual incision in the damaged-strand at the 5' and 3' of the lesion is carried out by XPF-ERCC1 and XPG, respectively. Each individually unstable (143,144), XPF (ERCC4) and ERCC1 stabilize each other by forming obligate heterodimers of XPF-ERCC1 (145) and perform the 5' incision with the same structure specificity towards 3' ssDNA overhangs as their *S. cerevisiae* homologs Rad1-Rad10 (146-149). In contrast, XPG (ERCC5) cleaves at the 3' junctions of 5' ssDNA overhangs (150). Remarkably, it has been shown that the 5' incision by XPF-ERCC1 is made first in the absence of the 3' incision by XPG, requiring only the presence of XPG at the complex (151). This 5' incision allows repair synthesis to begin, before XPG initiates the 3' incision, which is potentially activated by the arrival of the DNA polymerase (151). Such coordinated sequential incision mechanism allows repair synthesis to initiate from the 3'-OH resulting from the XPF-ERCC1 cleavage and minimizes the time, following the release of the TFIIH-bound damage-containing ssDNA (152), when a gapped (22-30 nt) DNA repair intermediate is present (153). Initial evidence suggested that either of the B family polymerases pol δ or pol ε could carry out repair synthesis in conjunction with the replication factor C(RFC) complex and proliferating cell nuclear antigen (PCNA) (154). It was later found that gap filling could involve three different polymerases recruited through different mechanisms: pol ε appears to be responsible for 50% of all repair synthesis, while pol δ and the Y family error-prone trans-lesion synthesis polymerase pol κ account for the rest (155,156). The same complexity is also involved in the last step of NER, ligation of the nick in DNA after repair synthesis, where either DNA ligase I or DNA ligase IIIa-XRCC1 is needed in a cell-cycle-dependent manner (157).

1.2 DISEASES ASSOCIATED WITH NER

1.2.1 Xeroderma Pigmentosum (XP)

Initially described by professor of dermatology Moriz Kaposi in 1874, xeroderma pigmentosum is a rare autosomal recessive disorder that affects approximately 1 in 1,000,000 people in the United States (158) and western Europe (159), and 1 in 22,000 in Japan (160). Skin of XP patients is extremely sensitive to UV exposure and develops hyper- and hypo-pigmentation in sun-exposed areas (161). Since the first piece of evidence linking XP to defective DNA repair presented by James Cleaver in 1968 (10), it has been established that there exist eight complementation groups of XP, named XP-A through XP-G and XP-V (161). These group names correspond to seven different NER gene products, XPA (162), XPB/ERCC3 (163,164), XPC (86), XPD/ERCC2 (165,166), DDB2 (XPE) (167), XPF/ERCC4 (145), XPG/ERCC5 (168), as well as polymerase η (XP-V) (169,170), which can correct (complement) the defective repair *in vivo* in add-back experiments. Loss of GG-NER can lead to an accumulation of lesions in the genome, resulting in increased mutagenesis.

A 2011 study of 106 XP patients examined at the NIH Clinical Center between January 1971 and December 2009 indicated that those under the age of 20 have 10,000-fold increase in non-melanoma skin cancer (NMSC) and 2,000-fold increase in melanomas (171). In addition to UV-exposed tissues, XP patients are also at higher risk for cancers in the brain and the central nervous system (172). Finally, 30% of XP patients also exhibit progressive neurodegeneration symptoms (173). The median age of death for XP patients was found to be between 29 and 37 years, depending on whether they had neurodegeneration (171). The 2011 study also showed that XP-C was the most common complementation group, accounting for 43% of the cohort (171).

Even though XP-C patients often do not exhibit extreme sun sensitivity or neurological abnormalities (174), they are more likely to develop skin cancers at an earlier age than patients in complementation groups XP-A, XP-B, XP-D, and XP-G (171). A 2016 study of 89 XP patients in the UK under long-term follow-up, of which 28 (31%) where diagnosed with XP-C, also identified a higher susceptibility of XP-C patients to severe ocular problems (175). With most of the disease-associated mutations found in the *XPC* gene resulting in protein truncations because of nonsense mutations, frameshifts, deletions, insertions, or splicing mutants, only five cases of missense mutations have been identified so far. They are: Pro334His in XP1MI (176), Tyr585Cys in XP107BR (175), Trp690Ser in XP13PV (177), Pro703Leu (178), and Thr738Ala (179).

1.2.2 Cockayne Syndrome, UV-Sensitive Syndrome, and Trichothiodystrophy

Cockayne syndrome (CS) is a neurodevelopmental disorder with a progeroid phenotype whose patients have an average life expectancy of just 12 years (180). While patients with CS are also UV-sensitive, unlike XP, their skin does not exhibit pigmentation changes after sun-exposure (181). Furthermore, CS patients do not appear to have an increased risk of cancer (180). Instead, symptoms of CS include retinopathy, microcephaly, demyelination, as well as retardation of growth and development (181). Mutations in CSA and CSB, which result in defective TC-NER, lead to a wide range of symptoms observed in CS patients (181). UV-sensitive syndrome ($UV^{S}S$) is another autosomal recessive disorder related to deficiencies in TC-NER (19). Only 8 patients who show mild symptoms in cutaneous photosensitivity have been identified with $UV^{S}S$ (31). Mutations in CSA (182), CSB (183), and UVSSA (29) genes have all been shown to give rise to $UV^{S}S$. Characterized by sulphur-deficient brittle hair and nail, patients with trichothiodystrophy (TTD) exhibit additional symptoms including developmental delay, short stature, ichthyosis, and

varying degrees of photosensitivity (184). Mutations in the p8/TTDA subunit of TFIIH have been identified in mildly photosensitive TTD-A patients (185), whereas mutations in the TTDN1 gene have been associated with non-photosensitive TTD patients (186,187). Finally, it is important to note that, due to their essential functions in both transcription initiation and NER damage verification, helicases XPB and XPD are often implicated in phenotypes combining symptoms of XP and CS or TTD (XP/CS) (181).

1.3 DAMAGE RECOGNITION BY XPC-RAD23B AND RAD4-RAD23

1.3.1 XPC-RAD23B and Rad4-Rad23

As reviewed above in **1.1.3**, Rad4-Rad23 is the *S. cerevisiae* homolog of the human XPC-RAD23B (60,86,87). Rad4 (754 a.a.) and Rad23 (398 a.a.) share 27% and 30% sequence identity with XPC (940 a.a.) and RAD23B (409 a.a.), respectively. Rad33, a functional homolog of CETN2, is also involved in NER (188) and interacts with Rad4 to form the Rad4-Rad23-Rad33 complex (189). Rad23 participates in NER by stabilizing Rad4, stimulating specific binding of the Rad4-Rad23 complex to damaged DNA, and targeting ubiquitinated proteins for proteasomal degradation (190,191). Cellular concentration of Rad23 ($\sim 1.1 \times 10^4$ copies/cell) is more than 10-fold excess over that of Rad4 (~ 880 copies/cell), suggesting that all soluble Rad4 molecules are likely in complex with Rad23 (192). In another parallel to XPC-RAD23B, Rad4-Rad23 has been shown to preferentially bind to DNA containing 6-4PPs, AAF adducts, fluorescein-modified bases, and bubbles (193-196). Finally, Rad4 is also known to recruit TFIIH through interactions with its Tfb1 (p62) and Ssl2 (XPB) subunits (197,198). The highly conserved functional features of these two

eukaryotic NER damage recognition complexes make Rad4-Rad23 an ideal subject of research on the topic of damage recognition, providing insightful results that are highly relevant to further understanding of the human XPC-RAD23B counterpart.

1.3.2 Molecular Mechanism of Damage Recognition

Solved in 2007, the crystal structure of Rad4-Rad23 in complex with a CPD-containing duplex DNA provides a detailed model for much of the research on damage recognition in the decade that followed (88). The overall structure of Rad4 (101 – 632) consists of four domains: a 310-residue transglutaminase-homology domain (TGD) and three β-hairpin domains (BHD1 – 3), each of which containing a long β-hairpin. TGD and BHD1 interact with the undamaged 11 bp dsDNA sequence through contacts made mostly to the DNA backbone. β-hairpins 2 and 3 in BHD2 and BHD3 approach the lesion from the minor and major groove, respectively. In particular, β-hairpin 3 inserts into the DNA duplex at the site of lesion, disrupting the double helical structure locally and causing the two mismatched bases opposite the CPD to flip out. These two thymidines are stabilized by the interactions with residues in both BHD2 and BHD3, whereas the CPD lesion becomes disordered. Furthermore, the DNA duplex appears kinked by approximately 42° in the region where β-hairpins 2 and 3 'pinch' on the lesion. Surprisingly, these aforementioned structural features of Rad4-Rad23 in a recognition complex with damage-containing DNA are replicated in a recent crystal structure where the protein is disulfide-tethered to an undamaged DNA duplex (199). Based on this observation and temperature-jump perturbation spectroscopy measurements of Rad4-induced lesion opening times, a kinetic gating model for damage recognition was proposed. The probability of forming a stable Rad4-Rad23-DNA recognition complex is suggested to be dependent on the interplay between the opening rate of DNA duplex and the residence time

of Rad4-Rad23 at the site (199). The kinetic gating model is further refined by detection of an early and β-hairpin-independent DNA twisting step (200). This brief (~100 – 500 μs) nonspecific DNA twisting is proposed to bridge the orders of magnitude difference between typical protein residence time per DNA base (~μs) and previously measured DNA lesion opening time (~ms), therefore allowing efficient interrogation of DNA by Rad4-Rad23 (200).

In addition to DNA conformational dynamics in damage recognition, atomic level insights regarding protein-DNA interactions that take place during this process have come from recent molecular dynamics (MD) simulations. Guided by the 2007 structures, restrained MD simulations revealed that as BHD2 and BHD3 approach the lesion, DNA bending and unwinding are accompanied by displacement of the CPD (201). Umbrella sampling approach was utilized in the simulations to break up the reaction coordinate into a series of smaller windows, such that sampling in regions beyond very high energy barriers can be improved. Three phenylalanine residues (F599, F597, and F556) in BHD3, previously identified to stack with and stabilize flipped-out bases (88), were shown to form a 'base flipping highway' for these mismatched thymidines, such that the base flipping occurs in a sequential fashion along the path defined by F599→F597→F556 (201). Finally, damage recognition (productive binding) is accomplished as β-hairpin 3 spontaneously inserts into the locally denatured lesion site, allowing proper stacking of F556 with one of the flipped-out bases and resulting in a final DNA bend angle of 47° (201). In contrast to the popular conformational capture mechanism (202) that would require a state in which either one or both bases are flipped out, this 'correlated motion' mechanism is associated with a much lower energy barrier for the protein-DNA complex to achieve productive binding (201). Identification of the 'correlated motion' mechanism naturally raises the question of whether recognition of other types of lesions also proceeds through the same steps. To that end, most recent MD studies of Rad4-

Rad23 binding to a benzo[α]pyrene adduct show that Rad4 recognizes the lesion by first capturing the 'pre-flipped' base opposite the adduct (203), suggesting that there exist multiple paths leading to robust damage recognition of structurally distinct lesions.

1.4 DIFFUSION[1]

Diffusive transport lies at the heart of a broad array of cellular processes. A specific topic of interest is how proteins perform diffusion, either one- or three-dimensional, in search of their targets in DNA. Such targets may be a particular DNA sequence in the case of a transcription factor, or a damaged base in the case of a DNA repair enzyme. This section begins with a brief review of the simple random diffusive process with a historical perspective, which is then followed by introduction of the diffusive search problem for DNA-binding proteins with cognate target sequences, *i.e.* facilitated diffusion and the speed-stability paradox.

1.4.1 Introduction to Diffusion

1.4.1.1 Brownian Motion

When observing pollen particles from the plant *Clarkia pulchella*, suspended in solution, through his single lens microscope in June of 1827, Scottish botanist Robert Brown noted their peculiar random jiggling motion (204). He went on to discover the same property of microscopic particles

[1] This section is adapted with permission from the following published manuscript : Kong, M. and Van Houten B. (2017) Rad4 recognition-at-a-distance: Physical Basis of conformation-specific anomalous diffusion of DNA repair proteins. *Progress in Biophysics and Molecular Biology*, **127**, 93-104.

suspended in liquids in other pollen grains, powders of fossil wood, window glass, minerals, rocks, and even a fragment of the Sphinx (204). In a follow up publication, Brown reiterated that such perplexing motion was exhibited by "extremely minute particles of solid matter, whether obtained from organic or inorganic substances, when suspended in pure water, or in some other aqueous fluids," and that it did not arise from currents in the fluid or as a result of evaporation (205). The random walk of microscopic particles in suspension has since been termed Brownian motion (Figure 1.5) in honor of Robert Brown.

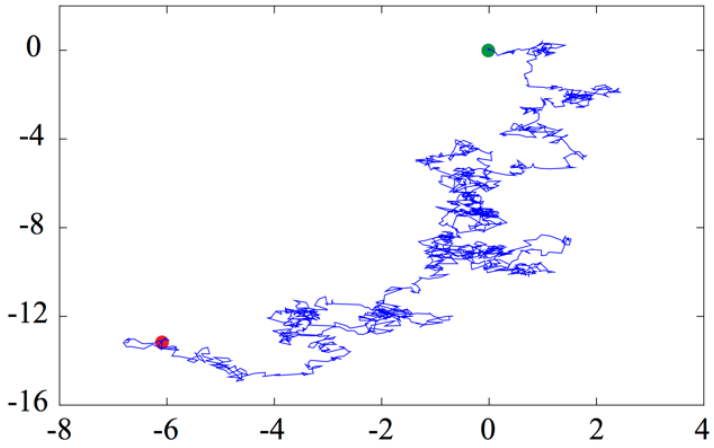

Figure 1.5: Simulated two-dimensional Brownian motion.

Green and red dots indicate the start and end of the trajectory, respectively. Adapted with permission from (206).

1.4.1.2 Fickian Diffusion

The first quantitative phenomenological description of macroscopic diffusion was developed by physiologist Adolf Fick in 1855, based on the idea of macroscopic concentrations and fluxes (207). Inspired by Fourier's law of heat conduction and Ohm's work on electric conductivity, Fick's first law proposes that the one-dimensional flux is inversely proportional to the concentration gradient:

$$j = -D\frac{\partial c}{\partial x} \tag{1.1}$$

where j is the flux in the units of number per unit area per unit time, c the concentration of particles in the units of number per unit volume, x in the units of length, and D the diffusion coefficient in the units of $length^2/time$. By invoking conservation of mass in combination with Fick's first law and the assumption that the diffusion coefficient D is a constant, we arrive at the law of diffusion in one dimension, or Fick's second law:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial^2 x} \tag{1.2}$$

Consider the case $c(x,t)$ where there the initial concentration at $t = 0$ is a spike at $x = 0$, or

$$c(x,0) = N\delta(x) \tag{1.3}$$

where $\delta(x)$ is the Dirac delta function (208). The solution to Fick's second law then takes the form

$$c(x,t) = \frac{N}{\sqrt{4\pi Dt}}e^{-x^2/4Dt} \tag{1.4}$$

i.e. a zero-mean Gaussian distribution that broadens over time (Figure 1.6). For a single particle, it can be shown that

$$\langle x^2(t)\rangle = 2Dt \tag{1.5}$$

This familiar result reflects the well-known linear relationship between time and the mean squared displacement (MSD) of a particle performing a one-dimensional random walk. This should not come as a surprise because the solution $c(x,t)$ is the probability distribution that characterizes the Wiener process, which in turn is the continuum limit of a one-dimensional random walk.

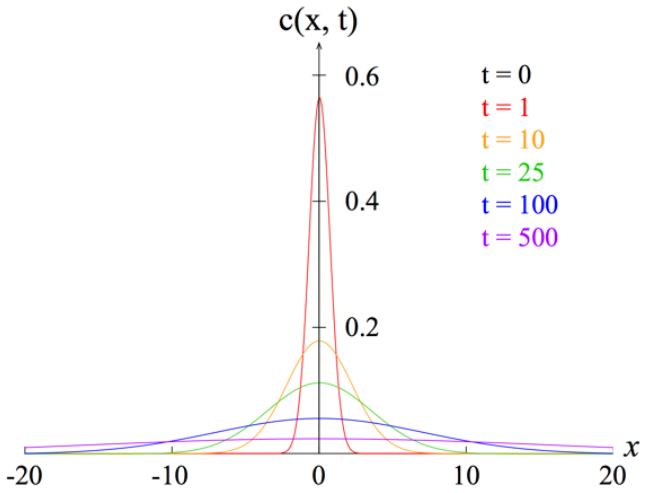

Figure 1.6: Time evolution of the solution to a one-dimensional Fickian diffusion that starts as a point source at the origin.

Adapted with permission from (206).

1.4.1.3 Einstein's Theory of Brownian Motion

In 1905, Albert Einstein took a more microscopic approach to the theory of diffusion (209,210). Focusing on the behavior of each independent diffusing particle suspended in a liquid, he arrived at the diffusion equation $\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial^2 x}$, its solution for the case of diffusion of n particles from a point source $f(x,t) = \frac{n}{\sqrt{4\pi Dt}}e^{-x^2/4Dt}$, as well as the linear dependence of the so-called mean squared displacement (MSD) on time $\langle x^2(t)\rangle = 2Dt$. These expressions are indeed equivalent to those from Fick's second law.

Another important finding presented in the same paper applied to the relation between the diffusion coefficient and other measurable physical quantities in diffusion. By considering a dynamic equilibrium where spherical particles were suspended in liquid and undergoing diffusion as a result of a force acting on them, Einstein derived the well-known Stokes-Einstein relation

$$D = \frac{k_B T}{6\pi\eta R} \tag{1.6}$$

where k_B is the Boltzmann constant, η is the viscosity of the fluid, and R is the radius of the diffusing particle (210). Its general form, $D = \mu k_B T$, where μ is the mobility of the particle, is also called the Einstein-Smoluchowski relation, as independently derived by Marian Smoluchowski in 1906 (211).

A more general approach to describe the motion of a Brownian particle of mass m is through the Langevin equation

$$m\frac{d^2x(t)}{dt^2} = -\gamma\frac{dx(t)}{dt} + \xi(t) \tag{1.7}$$

where $x(t)$ is the stochastic position of the particle, γ is the friction coefficient (the Stoke drag), and the stochastic term $\xi(t)$ represents a random fluctuating force on the particle (212). Such $\xi(t)$ has a Gaussian distribution with

$$\langle\xi(t)\rangle = 0, \langle\xi(t)\xi(t')\rangle = g\delta(t-t') \tag{1.8}$$

It can be shown that $\xi(t)dt = dW(t)$, where $dW(t)$ is a Weiner process by definition. At equilibrium, applying the equipartition theorem yields

$$g = 2\gamma k_B T \tag{1.9}$$

and

$$\langle(x(t) - x_0)^2\rangle = \frac{2k_B T}{\gamma}t = 2Dt \tag{1.10}$$

the Stokes-Einstein relation can be recovered

$$D = \frac{k_B T}{\gamma} = \frac{k_B T}{6\pi\eta R} \tag{1.11}$$

Coincidentally, in the same year that Einstein published his paper on the theory of diffusion, the term “random walk” was first used in a letter to *Nature* titled “The problem of the random walk,” by British statistician Karl Pearson (213). Pearson was originally interested in the spread of mosquito infestation and framed the problem as a man carrying out a random walk (Figure 1.5).

1.4.2 The Target Search Problem: Solving the Speed-Stability Paradox.

A simple question that has inspired biophysics research for the past half century is: how do limited copies of a sequence-specific DNA binding protein (*e.g.* the *lac* repressor, LacI, at ~10 molecules/cell) (214) efficiently locate its target that is buried in a sea of nonspecific sequence (~4.6×10^6 bp/cell) (215)? The answer appears to be a phenomenon called facilitated diffusion, as described below.

1.4.2.1 Facilitated Diffusion

The nature of diffusive transport of DNA binding proteins in the context of target search has been of intense interest for decades and has extensive implications in many different facets of essential cellular processes, ranging from DNA replication and gene regulation to maintenance of genome stability (216-219). The importance of search dimensionality was first pointed out by Adam and Delbruck, who suggested that the search process can be accelerated by collapsing a three-dimensional search into a one-dimensional search along the DNA (220). The theory was corroborated by the experimental observation that the association rate of the *lac* repressor to its target is two orders of magnitude faster than three-dimensional diffusion-based predictions according to its size and the viscosity of the media it travels through (221). The concept of *facilitated diffusion* was subsequently proposed and experimentally studied by von Hippel and colleagues, among others (222-224). In addition to three-dimensional diffusion in solution and one-dimensional sliding on DNA, the facilitated diffusion model also includes microscopic hopping of proteins on DNA, as well as direct intersegmental transfer of a protein between two DNA molecules (Figure 1.7).

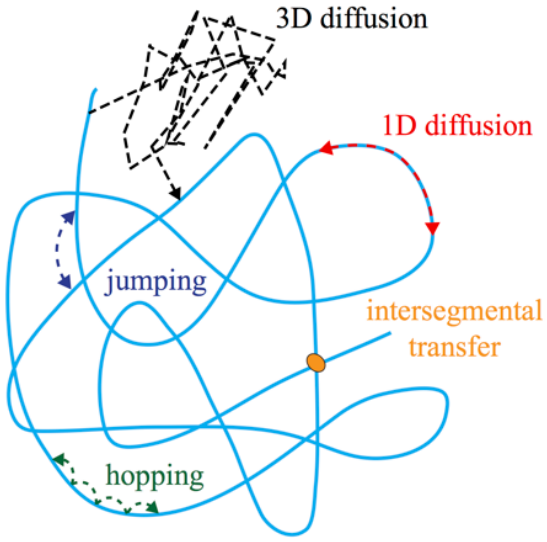

Figure 1.7: Schematic of facilitated diffusion.

Optimal search strategies combine different modes of protein-DNA interactions, including: 3D diffusion (black), 1D diffusion (red), jumping (blue), hopping (green), and intersegmental transfer. In the case of intersegmental transfer, the protein (orange) binds to two DNA molecules at the same time, releases from the one where it is initially bound, and transfers to the other molecule. Adapted with permission from (206).

Based on the frame work of facilitated diffusion, combining three-dimensional diffusion in solution and one-dimensional sliding on DNA, initial kinetic (225,226) and stochastic models (227) were established to address the optimal search strategy. As this field evolved, more recent studies combined other search modes, including hopping, jumping, and intersegmental transfer, with the effects of DNA conformation in their analyses (228-233). The effects of macromolecular crowding on facilitated diffusion, as it relates to more physiological conditions in living cells, have also been recently examined (234,235). The latest experimental (236) and theoretical studies (237-239) have shown that crowding environments can lead to altered balance between three-dimensional and one-dimensional diffusion processes, promoting one-dimensional sliding. While the presence of mobile or immobile obstacles on DNA has been shown to effectively slow down one-dimensional sliding (234,238,240), this effect could be overcome by hopping on DNA. Even

though emphasis in such studies is usually placed on the interplay between three-dimensional and one-dimensional search strategies, with three-dimensional diffusion being modeled as strictly Brownian, the potential for anomalous subdiffusion in crowded environments has nonetheless been noted in the context of facilitated diffusion (241,242).

Since the pioneering work by Riggs and coworkers (221), facilitated diffusion has been experimentally observed, both *in vitro* and *in vivo*, for a wide host of DNA-binding proteins. A short list of such proteins includes restriction enzymes (243,244), human transcription factor p53 (245), DNA repair proteins (219,246,247), and transcriptional repressors such as LacI (235,248-250).

1.4.2.2 Speed-Stability Paradox

While developing an optimization for target search, Slutsky and Mirny quantitatively formulated the speed-stability paradox of protein-DNA recognition (226). In brief, it was shown that rapid sliding of proteins on DNA with a sequence-dependent Gaussian-distributed energy landscape is only possible when the landscape is relatively smooth ($\sigma < 1 - 2\ k_B T$); however, conditions for stable binding necessitate a large variance in energy distribution ($\sigma > 5\ k_B T$), i.e. a rugged landscape (251). The authors proposed a two-state model as a solution to the paradox (Figure 1.8). The idea of a model based on protein conformational changes was first presented by von Hippel and colleagues (223). In summary, the protein, or protein-DNA complex in general, adopts two conformations: the recognition (R) state with a rugged energy landscape to allow stable binding and the search (S) state with a fairly smooth landscape to facilitate fast sliding. Such a model was supported by the experimental observation of structural flexibilities in dimeric *lac* repressor binding to specific and nonspecific DNA (252). Quantitatively similar results were also obtained by Hu and Shklovskii through a different approach investigating the effect of energy profile

disorder on the enhancement of search rates (253). More refined and generalized interpretations of the two-state model have also since been discussed (254-258).

The two-state model was elegantly applied in single-molecule studies of p53 searching for DNA binding sites, where it was shown that the C-terminus of the protein allows rapid sliding with a shallow energy surface, while the DNA binding domain interrogates the major groove for specific DNA sequences within a steep energy surface (259,260). Fitting the observed diffusion constants to a two-state model indicated that p53 would need to switch between conformations at a minimum rate of 10^3/s (259). Other examples that lend support to this model include proteins involved in mismatch repair (246,261) as well as transcription activator-like proteins (262).

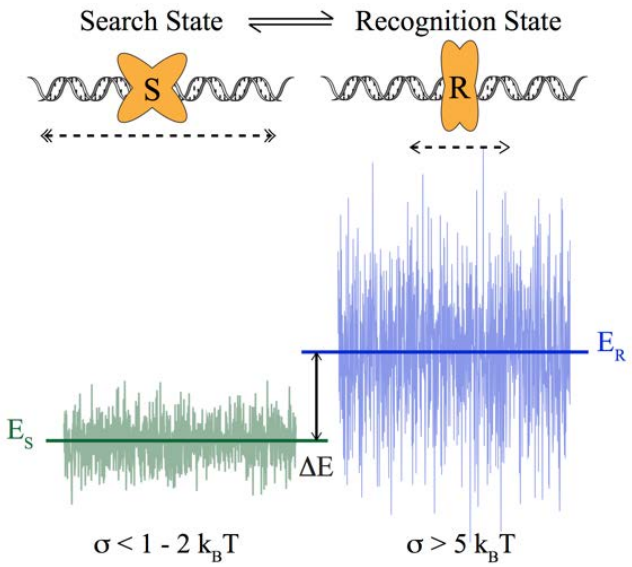

Figure 1.8: Schematic of the two-state model.

Protein (orange) is able to switch between two conformational states, *S* (search) and *R* (recognition). In the *S* state, protein slides fast (double arrow heads) on DNA with a smooth energy landscape. In the *R* state, protein slides slowly (single arrow heads) over a more rugged energy landscape. The equilibrium constant for transitions between states *S* and *R* depends on the energy difference ΔE between mean energies of the two states, E_S and E_R. Adapted with permission from (206).

1.5 HYPOTHESES AND SCOPE

In the years leading up to the beginning of this dissertation, significant progress was made towards the understanding of the molecular mechanism of damage recognition of XPC-RAD23B (Rad4-Rad23). Much of this advancement in the field was brought about in no small part by the 2007 Rad4-Rad23-DNA crystal structure from the Pavletich laboratory, which provided the molecular basis for the broad substrate specificity of this NER damage sensor. While *in vivo* dynamics involving XPC-RAD23B had also been examined using fluorescently-tagged protein constructs in the broader context of the entire NER process in response to UV-irradiation of cells (63,72,263,264), details that bridge the gap between these recent studies and years of *in vitro* biochemical work on XPC-Rad23B (Rad4-Rad23) had yet to emerge. During the past decade, the field of single-molecule biophysics had also enjoyed significant growth. Aided by the inventions of multiple techniques from the Greene, Kad, and van Oijen laboratories among others, the topic of protein target search on DNA, one that had been around since the 1960s, garnered tremendous attention thanks to the ability to now visualize each protein particle diffuse on long DNA substrates at the single-molecule level. Finally, to follow up on our earlier study of UV-DDB using defined lesion substrates (43) with the goal of reconstituting NER at the single-molecule level, we embarked on this study, in collaboration with Dr. Jung-Hyun Min at University of Illinois at Chicago, to address the following questions:

- *How does Rad4-Rad23 interact with DNA and conduct target search?* Based on findings from previous single-molecule studies on other DNA repair proteins, the hypothesis is that Rad4-Rad23 exhibits some combination of 3D and 1D diffusive behavior, *i.e.* facilitated diffusion. However, it remains to be seen how the process is partitioned into 3D vs. 1D schemes and exactly how 1D diffusion is carried out. Given that Rad4-Rad23 has been

biochemically characterized as to not bind to CPDs, yet is required *in vivo* for CPD removal, direct observation of its behavior on CPD-containing defined-lesion substrates is expected to shed more light on this situation.

- *What structural component(s) of Rad4 are important for damage recognition?* The 2007 crystal structure strongly suggests that β-hairpin 3 is the damage sensing motif of Rad4-Rad23. This hypothesis is tested through a series of structure-function relation studies, using both single-molecule and bulk biochemistry methods, with truncation mutants of Rad4 that lack various components of the BHD3. I predicted that the substrate specificity of Rad4-Rad23 resides in the β-hairpin 3. Loss of β-hairpin 3 in Rad4 renders it incapable of damage recognition and changes the diffusive behavior of the protein complex. An extension of this overarching question is whether any of the observed effects in *in vitro* experiments leads to a change in the response and survival of *S. cerevisae* after UV irradiation. Increased UV-sensitivity is expected to arise from cells carrying Rad4-Rad23 mutant(s) that have impaired damage sensing abilities.

The following chapters of this dissertation will detail the materials and methods used to conduct the proposed studies (Chapter 2.0), present the findings and their implications (Chapter 3.0), and further discuss the physical interpretation of these results, as well as prospects of future studies (Chapter 4.0).

2.0 MATERIALS AND METHODS

2.1 DEFINED-LESION SUBSTRATES

2.1.1 Incorporation of Site-Specific Lesion in pSCW01 Plasmid

Defined lesion substrates for single-molecule tightrope assays and atomic force microscopy (AFM) studies were constructed essentially as described before (43). pSCW01 plasmid was amplified using the *E. coli* DH5α strain overnight in Luria-Bertani liquid culture containing 100 µg/ml ampicillin. Plasmid DNA was extracted from harvested cells using the QIAGEN Plasmid Maxi Kit (QIAGEN, 12162) and concentrated to ~ 1 µg/µl. pSCW01 (>50 µg) was incubated with twice the number of units of the nickase Nt.BstNBI (New England BioLabs, R0607L) as the amount of DNA in micrograms and 50-fold molar excess of displacer oligonucleotides (Displacers 1, 2, and 3, Integrated DNA Technologies, Table 2.1) in 1X NEBuffer 3.1 (New England BioLabs, B7203S) at 55 ºC for 4 hours. The nicking reaction was then inactivated by heating at 80 ºC for 20 minutes and allowed to slowly cool down to room temperature while remaining in the heat block that had been turned off (~3.5 – 4 hours). During the cool-down process, the short oligonucleotides liberated from the nicking reaction were captured by the displacer strands. The now gapped plasmid was purified by addition of equal volume of the PEG solution (26% PEG-8000 in 20 mM $MgCl_2$) and centrifugation at 14,800 rpm for 1 hour at 4 ºC. Precipitated DNA was washed with

70% ethanol, pelleted by centrifugation at 14,800 rpm for 15 minutes at 4 ºC, air dried at room temperature, and finally re-suspended in autoclaved H_2O. To anneal the 37mer custom oligonucleotides containing the desired type of lesion (FL37, Integrated DNA Technologies or CPD37, TriLink, Table 2.1) to purified gapped plasmids, the plasmids were incubated with 3-fold molar excess of lesion-containing oligonucleotides in 1X NEBuffer 2.1 (New England Biolabs, B7202S) at 85 ºC for 10 minutes, before the heat block was turned off to allow the mixture to slowly cool down to room temperature (3.5 – 4 hours). To seal the nicks remaining in plasmids, the annealing reaction was supplemented with ATP (Thermo Scientific, R0411) and T4 DNA ligase (New England BioLabs, M0202M) to the final concentrations of 8 mM and 20 U/µl, respectively. The mixture was then incubated at 16 ºC for 18 hours, before the ligation reaction was inactivated by heating at 65 ºC for 10 minutes and allowed to cool down slowly to room temperature.

2.1.2 Substrates for DNA Tightrope Assay

To generate long DNA substrates suitable for DNA tightrope assay, lesion-containing pSCW01 plasmids were first digested with twice the number of units of XhoI (New England BioLabs, R0146L) as the amount of plasmids in micrograms, in 1X NEBuffer 2.1, at 37 ºC for 2 hours. XhoI was then inactivated by heating the digestion reaction at 65 ºC for 20 minutes, before the heat block was turned off and the mixture slowly cooled down to room temperature. For tandem (end-to-end) ligation, 1 µg of linearized plasmid DNA was incubated with 2 µl of T4 DNA ligase (New England BioLabs, M0202M) in 1X Quick Ligation buffer (New England BioLabs, M2200S) at room temperature for 15 minutes. The reaction tube was then immediately placed on dry ice to

stop the ligation reaction. Defined-lesion substrates therefore contained one site-specific lesion every plasmid length, or 2030 bp.

For some control experiments that required undamaged DNA, pSCW01 plasmids were not subjected to nicking, annealing, or ligation. Plasmid DNA was purified, linearized with XhoI, before tandemly ligated into long substrates.

2.1.3 Substrates for Atomic Force Microscopy (AFM)

To obtain DNA substrates of appropriate length (~500 bp) for AFM, lesion-containing pSCW01 was double digested with twice the number of units of XmnI (New England BioLabs, R0194L) and PciI (New England BioLabs, R0655L) as the amount of DNA in micrograms, in 1X NEBuffer 2.1, at 37 ºC for 4 hours. The digestion reaction was then inactivated by heating the mixture at 80 ºC for 20 minutes, before turning off the heat block to allow it to slowly cool down to room temperature. The 538 bp lesion-containing fragment was separated on a 1% agarose gel, extracted using the Wizard SV Gel and PCR Clean-Up System (Promega, A9281), and purified again with the QIAquick PCR Purification Kit (QIAGEN, 28104) to remove restriction enzymes. Final elution of purified 538-bp DNA was made in autoclaved and 0.02 µm-filtered H_2O (GE Healthcare Life Sciences, Whatman Anotop Syringe Filters, 0.02 µm). Standard phenol-chloroform-isoamyl alcohol extraction followed by ethanol precipitation was also performed if proteins were found to remain DNA-bound upon initial quality check on AFM. The final 538 bp DNA duplex contained a site-specific lesion at 30% contour length from one end of the DNA.

2.2 UV-IRRADIATION OF λ-DNA

Commercially available λ-DNA (New England BioLabs, N3013L) was first diluted in 1X TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) to the final concentration of 50 ng/µl before aliquoted onto a weighing dish as individual drops of 20 µl in volume. UV-irradiation was carried out by placing the weighing dish under a germicidal UV lamp (254 nm) for the appropriate amount of time, as calibrated by a UVX Digital Radiometer with UVX-25 sensor (UVP), for the desired dose (20 or 40 J/m^2). Relationship between UV dose and lesion frequency had been previously characterized using qPCR (43,265). At 20 J/m^2 and 40 J/m^2, λ-DNA contained on average about 1 lesion (CPD or 6-4PP) per 2000 bp and 1000 bp, respectively (43), distributed randomly.

2.3 SINGLE-MOLECULE DNA TIGHTROPE ASSAY

Single-molecule DNA tightrope assay was performed as described previously (43,219). Briefly, No. 1½ coverslips ($24 \times 40\ mm$, Corning) were cleaned and treated with PEG (Laysan Bio Inc., mPEG-SVA MW 5,000) solution overnight. Custom flow cells were assembled by attaching the coverslips, via double-sided tape spacers, to microscope slides with glued inlet and outlet tubing. Prior to setting up DNA tightropes, flow cells were passivated by incubating with blocking buffer (50 mM HEPES pH 7.5, 50 mM NaCl, 1 mg/ml bovine serum albumin) for at least 10 minutes. This blocking step was followed up with pipetting 5 µm silica beads (Polysciences Inc., 24332-15), coated with poly-L-lysine (Wako Chemicals USA, 163-19091), into the flow cell. Randomly distributed beads were allowed to settle for 10 minutes before the flow cell was washed and connected to a 5 ml glass syringe (Hamilton, 81520) mounted on a syringe pump (World Precision

Instruments). DNA tightropes were formed by flowing 1 – 2 µg of DNA, diluted to a final volume of 100 µl in TR buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 3 mM $MgCl_2$), back and forth through the flow cell at the rate of 0.3 ml/minute for 45 minutes to 1 hour. If tandem-ligated defined-lesion substrates were used, following DNA string-up, the flow cell was first washed with 200 µl of high salt buffer (20 mM HEPES pH 7.5, 1 M NaCl, 3 mM $MgCl_2$), followed by equilibration with 400 µl of Rad4-Rad23 binding buffer (5mM BTP-HCl pH 6.8, 75 mM NaCl, 5% glycerol, 0.74 mM CHAPS, 0.5 mg/ml BSA, and 5 mM DTT).

Prior to imaging, purified Rad4-Rad23 with a N-terminal 6X-Histidine tag on Rad4 was labeled with 655 nm streptavidin-coated quantum dots (Invitrogen, Q10123MP) via Penta-His antibody biotin conjugate (QIAGEN, 34440). Specifically, 1 µl of 1 µM streptavidin-coated quantum dots were incubated with 5 µl of 1 µM biotin-conjugated antibody at 4 ºC for 1 hour. 1 µl of this quantum dot-antibody complex (~167 nM) was then incubated with 1 µl of His-tagged Rad4-Rad23 at 4 ºC for 1 hour. Finally, 2 µl of the quantum dot-labeled Rad4-Rad23 (~83 nM) was diluted in binding buffer to a final volume of 100 µl and injected into the flow cell at the final concentration of ~1.6 nM. All binding experiments, unless otherwise noted, were carried out in this binding buffer containing 5mM BTP-HCl pH 6.8, 75 mM NaCl, 5% glycerol, 0.74 mM CHAPS, 0.5 mg/ml BSA, and 5 mM DTT.

Oblique angle fluorescence imaging was performed using a Nikon Eclipse Ti inverted microscope base with a Nikon 100X TIRF objective with 1.45 numerical aperture. Quantum dots (QDs) were excited using a 488-nm laser with a power of 1-2 mW at the back focal plane. QD emissions were captured on an Andor Neo sCMOS camera using Nikon Elements Ar software and saved as ND2 files. Exposure time was either 80 or 100 ms per frame, which resulted in frame rates of 10.92 or 10 fps, respectively. Particles were typically recorded for five minutes, and the

proteins refreshed every two hours. The positional accuracy and localization precision of the system have been reported previously to be ~6 nm and ~10 nm, respectively (43).

For data analysis, each ND2 file was first exported as a time series of TIFF images using Nikon Elements Viewer, imported into ImageJ software (NIH), and saved as a TIFF stack. Kymograph of each protein particle of interest was generated by using the slice function in ImageJ on the image stack, over a line drawn along the trajectory of the particle. Position of the particle was then tracked by fitting the intensity profile of the kymograph at each time point with a one-dimensional Gaussian in ImageJ. Positional data was transferred to Matlab (MathWorks), where given the pixel size (43 nm) and frame rate, a custom-written script calculated the mean squared displacement (MSD). To obtain diffusion coefficient D and anomalous diffusion exponent α, the MSD vs. time step (Δt) curve was fitted to the general model of 1D diffusion (266)

$$MSD = 2D(\Delta t)^{\alpha} \tag{2.1}$$

The minimum criteria for reporting fitted D and α values were such that at least 10% of the entire diffusion trajectory was used in a fitting that produced an R^2 value of 0.8 or higher.

Several criteria were employed consistently to ensure unbiased classification of particle motion. First, a particle was considered to be motile if its Gaussian-fitted position displacement was greater than three pixels, or ~400 bp at 43 nm/pixel, whereas motions below three pixels were considered nonmotile. Even though this criterion did act as a constraint in scoring motile particles, as pointed out earlier, the positional accuracy was much higher (~30 bp). As a result, constrained motion may have been conservatively underestimated. To differentiate random versus constrained motion, the criterion arose from the examination of kymographs: random motion covered several thousand base-pairs before reversing direction and usually the total track length of these particles was great distances (10-30 kbp) along the DNA tight rope; whereas constrained motion was seen

an oscillatory motion around a central point, the motion in each direction being usually 500-1000 bp.

2.4 ATOMIC FORCE MICROSCOPY

600 nM Rad4-Rad23 was incubated with 150 nM Fl-dT-containing (30% contour length from one end) DNA fragments at room temperature for 30 minutes in 0.02 μm-filtered AFM binding buffer (5mM BTP-HCl pH 6.8, 75 mM NaCl, 5% glycerol, and 0.74 mM CHAPS). Binding reaction was then diluted 50-fold in AFM deposition buffer (25 mM HEPES pH 7.5, 25 mM NaOAc, and 10 mM $Mg(OAc)_2$). 10 μl of the dilution was immediately deposited onto freshly cleaved mica surface (SPI Supplies, 1877-MB) for 30 seconds, rinsed with 1 ml of 0.02 μm-filtered H_2O, and dried in a gentle stream of compressed nitrogen. AFM data were collected on a MultiMode V microscope (Bruker Corp.) with an E scanner and Pointprobe® plus noncontact silicon probes (PPP-NCL, Agilent Technologies) in tapping mode. Images (512 × 512 pixels) were captured at a scan rate of ~3 Hz over areas of 1 μm × 1 μm in size using NanoScope 7 software (Bruker) and flattened. For bend angle analysis, AFM images were first exported as TIFF files with NanoScope 7 and then processed with the built-in segmented line and angle tools of ImageJ. Only DNA molecules not crossing itself or others, with both ends clearly visible, and internally-bound protein present were analyzed. Specifically, the contour of each candidate DNA molecule was traced using the segmented line tool in ImageJ. The total length in pixels was converted to nm using the conversion factor of 1000 nm = 512 pixels. Protein binding positions and total DNA contour lengths were measured in pixels and binding positions were then converted to percentage DNA contour length from the closest end, and thus always ≤ 50%. For each internal binding event, the angle tool was

used to quantify the degree of DNA bending by manually tracing the paths of DNA immediately exiting through the two sides of protein particles. This measured angle was always the smaller angle of the two and therefore less than or equal to 180°. DNA bend angle was then obtained by subtracting the measured angle from 180° to reflect bending from a straight DNA molecule. Intrinsic DNA bend angles at Fl-dT sites were quantified by first determining positions that are 30% from either end of each DNA molecule with the segmented line tool, and then measuring bending with the angle tool at these positions for all molecules.

2.5 CPD AND 6-4PP REPAIR KINETICS BY ANTIBODY SLOT BLOT

S. Cerevisiae carrying WT and mutant Rad4 constructs were first grown in YPD media overnight. The amount of overnight culture needed was calculated and added to 50 ml YPD media such that the new culture had an initial reading of OD600 ~ 0.1. The culture was then grown in the dark at 30 ºC for ~5 hours, shaking at 220 rpm, to log phase (OD600 ~ 0.8). Cells were harvested via centrifugation at $1{,}000 \times g$ for 5 minutes, re-suspended in 1X PBS in twice the original volume, transferred to plastic petri dishes in aliquots of 10 ml, and irradiated with 100 J/m^2 UVC in the dark. For all steps carried out after UV irradiation, cells were shielded from light to minimize photoreactivation. Irradiated cells were first pooled, placed on ice temporarily, before collected by centrifugation at $1{,}000 \times g$ for 5 minutes and re-suspended in the same volume of YPD media without dilution. For the zero time point, an aliquot of 9 ml was taken, collected via centrifugation at $2{,}000 \times g$ for 5 minutes, and flash frozen in a mixture of 100% ethanol and dry ice. The remaining cells in YPD were allowed to recover in the dark at 30 ºC for up to five hours, shaking at 220 rpm. At time points 0.5, 1, 2, and 5 hours after beginning of dark recovery, aliquots of 9 ml

were taken from the culture, collected via centrifugation at $2{,}000 \times g$ for 5 minutes, and flash frozen. Genomic DNA from 9-ml cell culture aliquots taken at various time points was extracted with YeaStar Genomic DNA Kit (Zymo Research, D2002) and quantified with PicoGreen (Invitrogen, P7581) on a NanoDrop 3300 fluorospectrometer (Thermo Scientific, ND-3300).

Antibody slot blot assays were performed as previously described (267). Genomic DNA samples from each time point, along with 2-fold serial dilutions of 40 J/m^2 UV-irradiated λ-DNA as standards, were diluted 2-fold with equal volume of 2X SSC buffer (Sigma, S6639), heated at 95 ºC for 5 minutes, and immediately placed on ice. For each recovery time point, genomic DNA samples were spotted in triplicates on Amersham Hybond-N+ membrane (GE Healthcare Life Sciences, RPN303B) with Minifold I Spot-Blot System (GE Healthcare Life Sciences, 10447850). Dilutions of UV-irradiated λ-DNA standards were spotted only once. Each spot contained 20 ng of genomic DNA, or 2-fold dilutions of UV-irradiated λ-DNA starting at 20 ng, in 100 µl volume. Spotted DNA samples on membrane were incubated with denaturing solution (1.5 M NaCl and 0.5 N NaOH) for 8 minutes, followed by neutralizing solution (0.5 M Tris-HCl pH 7.0 and 1M NaCl) for 8 minutes. Membrane was then baked, in a sandwich between Whatman filter paper (GE Healthcare Life Sciences, 10447850), at 80 ºC for 2.5 hours, and blocked at room temperature in 1X TBS containing 0.3% Tween20 (TBS-T) and 5% non-fat dry milk for one hour. UV lesions were probed by incubating membrane overnight with primary mouse antibodies against CPD (1:2,000, Kamiya Biomedical, MC-068) or 6-4pp (1:2,000, Cosmo Bio, CAC-NM-DND-002) in 1X TBS-T at 4 ºC. Blots were washed in 1X TBS-T, incubated with HRP-conjugated anti-mouse secondary antibody for 1 hour at room temperature when appropriate, developed with Supersignal West Femto Maximum Sensitivity Substrate (Thermo Scientific, 34095), imaged on ChemiDoc

MP (Bio-Rad), and quantified using Image Lab software (Bio-Rad). Under these experimental conditions, 100 J/m^2 produced ~2.6 CPD and ~1 6-4PP per 2 kbp.

2.6 AGAROSE GEL ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA)

1 µM Rad4-Rad23 was incubated with equal amount of 1 µM biotinylated His-Antibody (QIAGEN, 34440) at 4 ºC for 1 hour. 500 nM Rad4-HisAb complex was then incubated with equal amount of 1 µM 705 nm streptavidin-coated quantum dots (Invitrogen, Q1063MP) at 4 ºC for 1 hour, shielded from light. In the HisAb-QD control, 500 nM HisAb was incubated with equal amount of 1 µM 705 nm QD at 4 ºC for 1 hour, shielded from light. Binding reactions were carried out in 10 µl volumes at room temperature for 30 minutes with 10 nM 37 bp Fl-dT-containing dsDNA (FL37-RC37, Integrated DNA Technologies, Table 2.1) and one of the following: 200 nM Rad4-Rad23, 200 nM Rad4-HisAb, 167 nM Rad4-HisAb-QD, 200 nM HisAb-QD, or 200 nM QD. A thin 1% agarose gel was cast in ½X TBE and pre-run on ice at 80V for 30 minutes before samples were loaded and run on ice at 80V for 60 minutes. Gel image was scanned using a Typhoon 9400 Variable Mode Imager (GE Healthcare Life Sciences) set to fluorescent mode with excitation wavelength 488 nm, emission wavelengths 526 nm and 670 nm for detection of fluorescein and QD signals, respectively.

2.7 FLUORESCENCE ANISOTROPY

Saturation binding experiments of increasing concentrations of Rad4-Rad23 to 7.7 nM of 50bp Fl-dT-containing duplex DNA fragment (26F50-NDB50, Integrated DNA Technologies, Table 2.1) were performed on a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies) at 37 °C using a set of four matched quartz cuvettes in binding buffer. Embedded fluorescein was excited at 485 nm and its emission collected at 520 nm, with slit widths of 5 and 10 nm for excitation and emission, respectively. The G-factor, a sample-independent correction factor for polarization bias of the fluorescence detection system, was determined once at the beginning of each experiment. At each titration point, protein was added and the cuvette tapped for mixing. After 3 minutes of incubation, anisotropy measurements were made in triplicates. Final reading was reported as mean ± standard deviation. At the end of each experiment, 5M NaCl was added to cuvette to reach 1M NaCl final concentration. All binding events were shown to be reversible as anisotropy returned to initial values upon the addition of 1 M NaCl. To obtain equilibrium dissociation constants (K_d), binding isotherms (anisotropy vs. protein concentration) were fitted to the following single-site specific binding with Hill slope model using *Prism 6* (GraphPad Software):

$$r = a \cdot \frac{x^n}{{K_d}^n + x^n} + b \tag{2.2}$$

where r is anisotropy, a is maximum anisotropy at saturation, x is protein concentration, n is the Hill coefficient, K_d is the equilibrium dissociation constant, and b is the initial anisotropy without protein (268). Parameters n and K_d were shared across repeats of the same protein variant for global fitting.

2.8 ESTIMATION OF HYDRODYNAMIC RADII

Radius of gyration ($R_{g,Rad4} = 3.03\ nm$) of WT Rad4-Rad23 was estimated by WinHYDROPRO (269). Crystal structure of the protein complex (2QSF) was imported into the software and the following constants were used for estimation with the 'shell-model from atomic level' option: molecular weight $MW = 96.94\ kDa$, room temperature $\mathrm{T} = 298\ \mathrm{K}$, viscosity of water $\eta = 0.89 \times 10^{-3}\ \mathrm{Pa \cdot s}$ at room temperature (298 K). Hydrodynamic radius (R_H) was then estimated through the relationship:

$$R_{H,Rad4} = \frac{R_{g,Rad4}}{\rho} = 3.91\ nm \tag{2.3}$$

where $\rho = (\frac{3}{5})^{1/2}$ for spherical molecules (270).

An effective hydrodynamic radius of the quantum dot-labeled Rad4-Rad23 complex was estimated based on the combined hydrodynamic volume of a quantum dot (QD) and Rad4-Rad23 such that:

$$R_{H,eff} = ({R_{H,QD655}}^3 + {R_{H,Rad4}}^3)^{1/3} = 11.65\ nm \tag{2.4}$$

where $R_{H,QD655} = 11.5\ nm$ was the hydrodynamic radius of a 655 nm SAQD (271).

2.9 ESTIMATION OF THEORETICAL LIMIT OF DIFFUSION COEFFICIENT

The diffusion coefficient *D* for a sphere diffusing in fluid can be calculated using the Stokes-Einstein equation:

$$D = \frac{k_B T}{\xi} \tag{2.5}$$

where ξ is the friction term, k_B is the Boltzmann constant and T is temperature. At room temperature (T $=$ 298 K), $k_BT = 4.11 \times 10^{-21}J$. For a globular protein of radius R diffusing linearly on DNA while tracking the helix in solution with viscosity η, the friction term ξ has been derived (272):

$$\xi = 6\pi\eta R + \left(\frac{2\pi}{10.5\ BP}\right)^2 [8\pi\eta R^3 + 6\pi\eta R_{OC}(R)^2] \tag{2.6}$$

where $1\ BP = 0.34\ nm$ and R_{OC} is the off-center distance from the center of mass of the protein to the helical axis of DNA. Using the effective hydrodynamic radius $R_{H,eff} = R$ as calculated above, under the idealization that the sphere was offset from the DNA helical axis by the radius of DNA ($r_{DNA} = 1\ nm$) such that $R_{OC} = R_{H,eff} + r_{DNA} = 12.65\ nm$, we obtained the theoretical limit to diffusion for the QD-Rad4 complex:

$$D_{lim} = \frac{k_BT}{6\pi\eta R + \left(\frac{2\pi}{10.5\ BP}\right)^2 [8\pi\eta R^3 + 6\pi\eta R_{OC}(R)^2]} = 0.0199\ \mu m^2/s \tag{2.7}$$

2.10 CALCULATION OF ENERGY BARRIER TO FREE DIFFUSION

The energy barrier to free diffusion can be calculated by using the measured and theoretical diffusion coefficients D and the Arrhenius relation (219,247):

$$k = e^{-E/k_BT} \tag{2.8}$$

where k is the stepping rate in units of *steps/s* and related to D such that

$$k = 2D/l^2 \tag{2.9}$$

where l is the step size, assumed to be a single base pair. Therefore, the barrier to free diffusion is:

$$\Delta E = \ln\left(\frac{k_{lim}}{k_{expt}}\right) = \ln\left(\frac{D_{lim}}{D_{expt}}\right) \cdot k_BT \tag{2.10}$$

where D_{lim} is the theoretical limit to diffusion coefficient as calculated above, and D_{expt} is the experimental measurement. Using the equation above and experimental values for diffusion coefficients of constrained and random Rad4-Rad23 WT on UV irradiated λ-DNA, $0.0040 \pm 0.0013\ \mu m^2/s$ and $0.0138 \pm 0.0033\ \mu m^2/s$ respectively, we calculated that energy barriers to free diffusion for constrained and random movers were

$$\Delta E_{cons} = 1.60 \pm 0.32\ k_B T \tag{2.11}$$

$$\Delta E_{rand} = 0.37 \pm 0.24\ k_B T \tag{2.12}$$

2.11 ESTIMATION OF RESIDENCE TIME AT EACH BASE PAIR

The dwell time (τ_{bp}) of Rad4-Rad23 at each base pair during linear diffusion was estimated as the inverse of the stepping rate

$$\tau_{bp} = \frac{1}{k} = \frac{1}{2D/l^2} = l^2/2D \tag{2.13}$$

For constrained motion, $\tau_{bp} \cong 15\ \mu s$. For random diffusion, $\tau_{bp} \cong 4\ \mu s$. This microsecond time scale of Rad4 residence time per site was largely consistent with that of various other proteins undergoing one-dimensional diffusion on DNA and indeed much slower than the expected undamaged DNA opening times of Rad4, estimated to be much longer than 5-10 ms (199).

2.12 ESTIMATION OF MINIMUM TARGET SITE ENERGY

Given the genome size M, the minimum energy requirement at target site (E_0) for efficient recognition is given by (226)

$$E_0 = \sqrt{2 \ln M}\ k_B T \tag{2.14}$$

For yeast genome $M = 1.21 \times 10^7 bp$ (273), $E_0 = 5.71\ k_B T$.

2.13 ESTIMATION OF GENOME SEARCH TIME

Diffusion coefficients of unlabeled Rad4-Rad23 can be predicted using the Arrhenius equation and the energy barrier calculations from above (266). Specifically:

$$D_{Rad4,cons} = D_{lim} \cdot e^{-\frac{\Delta E_{cons}}{k_B T}} = 0.091 \pm 0.030\ \mu m^2/s \tag{2.15}$$

$$D_{Rad4,rand} = D_{lim} \cdot e^{-\frac{\Delta E_{rand}}{k_B T}} = 0.312 \pm 0.076\ \mu m^2/s \tag{2.16}$$

Whereas the search range ($\mathcal{R}$) for each encounter of Rad4 with DNA can also be estimated as $\mathcal{R} = \sqrt{\frac{16Dt}{\pi l^2}}$ where l is a single base pair and t is the time of encounter (274), or lifetime of Rad4-Rad23 on DNA (Figure B.5).

$$\mathcal{R}_{Rad4,cons} = (5.5 \pm 1.0) \times 10^3 bp \tag{2.17}$$

$$\mathcal{R}_{Rad4,rand} = (2.79 \pm 0.45) \times 10^4 bp \tag{2.18}$$

Knowing the yeast genome size and that there are ~870 copies of Rad4 per cell (192), we can calculate the least number of encounters required such that the total diffusion range covers the entire genome (266). Assuming distribution of motion types similar to that observed from WT protein on UV irradiated λ-DNA:

$$\#\ encounters = \frac{1.2 \times 10^7 bp}{\#\ Rad4 \cdot (\%constrained \times \mathcal{R}_{Rad4,cons} + \%random \times \mathcal{R}_{Rad4,rand})} = 2.22 \pm 0.29 \tag{2.19}$$

Therefore, the fastest possible time needed to search the entire genome of yeast is

$$time\ required = \#\ encounters \times \frac{time}{encounter} = 145 \pm 35\ sec \tag{2.20}$$

or roughly two to three minutes.

2.14 *RAD4* MUTANT STRAIN CONSTRUCTION, UV SURVIVAL MEASUREMENTS, AND WESTERN BLOTTING

RAD4 deletion mutants were generated in yeast using the CRISPR-Cas9 system, as described previously (275). Briefly, wild-type yeast strains (*RAD4* or *RAD4-3×FLAG*) were transformed with two plasmids, pT022 and pT040-RAD4, to generate double strand breaks at the *RAD4* gene that were repaired using oligonucleotide templates containing the desired Rad4 deletions (e.g., *rad4 Δ599-605*). pT022 expresses the Cas9 endonuclease and contains the *LEU2* selection marker. pT040-RAD4 harbors the *URA3* selection marker and expresses a single guide-RNA (sgRNA) that targeted the Cas9 endonuclease to the yeast *RAD4* gene to generate a DNA double strand break. By providing donor DNA (double stranded oligonucleotides, see Table B.2) containing Rad4 deletions, the targeted *rad4* mutations were introduced into the chromosomal *RAD4* locus by homologous recombination. *RAD4* deletion mutants were confirmed by PCR amplification of isolated genomic DNA and verified by DNA sequencing. The pT022 and pT040-RAD4 plasmids were removed from the *rad4* mutant strains by screening on plates lacking Leucine and 5-Fluoroorotic acid (FOA) counter selection.

To measure UV sensitivity, yeast cells were grown to mid-log phase in YPD ($OD_{600} \approx 0.8$). Each culture was serially diluted, spotted on YPD plates, and irradiated with UV light (254 nm) at the indicated UV doses. Plates were immediately wrapped with aluminum foil after UV treatment and incubated at 30˚C for ~72 hr. For the quantitative UV survival assay, yeast cells were spread

on YPD plates and then irradiated with UV light at the indicated doses. Plates were incubated in the dark at 30˚C prior to colony counting to determine the percentage of viable colonies.

To assess Rad4 protein level in each yeast strain, yeast whole cell extracts were prepared from 5 ml of yeast culture. Cells are incubated in 0.1 M NaOH for 5 min at room temperature, followed by boiling cells in 100 µl of 1× SDS-PAGE sample buffer for 7 min. After centrifugation, different volumes (3, 6, and 9 µl) of the supernatant for each strain were loaded in an 8% SDS gel. The presence of FLAG-tagged Rad4 protein was detected using anti-FLAG antibody (M2, Sigma). Tubulin was also probed with anti-tubulin antibody (Santa Cruz Biotech) as the loading control.

2.15 CPD REPAIR KINETICS BY T4 ENDO V DIGESTION

Exponentially growing yeast cells carrying WT and mutant Rad4 constructs were irradiated with 50 J/m^2 UVC and then repaired for 60 or 120 minutes. Genomic DNA was isolated and digested with CPD-specific T4 endo V to generate single strand breaks at CPD sites. DNA was then separated on a 1.2% alkaline agarose gel and fragmented DNA was detected with SYBR Gold (276). DNA signal was scanned with a Typhoon FLA 7000 laser scanner (GE Healthcare Life Sciences), and repair was analyzed using ImageQuant software (GE Healthcare Life Sciences). Under these experimental conditions, 50 J/m^2 produced 1 CPD per 4 kbp.

2.16 ESTIMATION OF RATE OF PHOTOPRODUCT REMOVAL

Rates of photoproduct removal by WT Rad4-Rad23 can be calculated from estimated lesion frequency (1 CPD/4 kbp under 50 J/m^2 in T4 Endo V digestion assays, ~2.6 CPD and ~1 6-4PP per 2 kbp under 100 J/m^2 in antibody slot blot assays) based on the copy number of Rad4-Rad23 (~870/cell) (192) and $T_{1/2}$ (times to accomplish 50% repair).

$$Rate\ of\ Photoproduct\ Removal = \frac{Lesion\ Frequency \times yeast\ genome\ size}{2 \times Rad4\ copy\ numer \times T_{1/2}} \quad (2.21)$$

From T4 Endo V digestion assays

$$Rate\ of\ CPD\ Removal \sim 1.6\ CPDs/Rad4/hour \quad (2.22)$$

From antibody slot blot assays:

$$Rate\ of\ CPD\ Removal \sim 3\ CPDs/Rad4/hour \quad (2.23)$$

$$Rate\ of\ 64PP\ Removal \sim 3.6\ 64PPs/Rad4/hour \quad (2.24)$$

Name	Sequence	Vendor
Displacer 1	ATTTGACTCC	IDT
Displacer 2	CATGGACTCGCTGCAG	IDT
Displacer 3	GAATGACTCGG	IDT
FL37	CCGAGTCATTCC**T**GCAGCGAGTCCATGGGAGTCAAAT	IDT
CPD37	CCGAGTCA**TxT**CCTGCAGCGAGTCCATGGGAGTCAAAT	TriLink
RC37	ATTTGACTCCCATGGACTCGCTGCAGGAATGACTCGG	IDT
26F50	GACTACGTACTGTTACGGCTCCATC**T**CTACCGCAATCAGGCCAGATCTGC	IDT
NDB50	GCAGATCTGGCCTGATTGCGGTAGAGATGGAGCCGTAACAGTACGTAGTC	IDT

Table 2.1: Oligonucleotides used for constructing defined-lesion substrates, as well as performing EMSA and fluorescence anisotropy experiments.

T = fluorescein-attached to the 5-position of the thymine ring with a 6-carbon spacer

TxT = Cys-Syn thymidine dimer

3.0 RESULTS[2]

The *Saccharomyces cerevisiae* Rad4-Rad23 complex is homologous to the human XPC-RAD23B, sharing structural and functional similarities (60,86). An X-ray crystal structure of Rad4-Rad23 bound to a 24-base pair duplex DNA harboring a CPD-containing mismatch lesion reveals that Rad4 consists of a transglutaminase homology domain (TGD) and three β-hairpin domains (BHD 1 – 3, Figure 3.1) (88). β-hairpin 3 of BHD3 (hereafter β-hairpin 3) is inserted into the DNA double helix at the lesion site. While the BHD2-BHD3 groove holds the two nucleotides displaced from the undamaged strand opposite the lesion, the CPD is flipped out of the helix and away from the protein. Finally the DNA duplex shows a kink of ~42°, similar to that resulting from binding of XPC-RAD23B to a cholesterol moiety in DNA, as measured by scanning force microscopy (277). Because of the lack of direct contact with the lesion by Rad4, Rad4 (XPC) is proposed to indirectly recognize locally destabilized duplex DNA by probing the two strands' propensity to open, which allows insertion of β-hairpin 3 (88). This hypothesis provides a working model for how Rad4 (XPC) recognizes chemically and structurally diverse DNA damage *in vitro*, such as a cholesterol-modified nucleotide, 6-4PP, cisplatin 1,3-d(GTG) intrastrand adduct, C8-dG acetylaminofluorene, and 5R-thymine glycol (76,78,82,84,193,278). Previous studies on domain deletions and mutated

[2] This chapter is adapted with permission from the following published manuscript: Kong, M., Liu, L., Chen, X., Driscoll K.I., Mao, P., Böhm, S., Kad, N.M., Watkins, S.C., Bernstein, K.A., Wyrick, J.J., Min, J.H., and Van Houten, B. (2016) Single-Molecule Imaging Reveals that Rad4 Employs a Dynamic DNA Damage Recognition Process. *Molecular Cell*, **64**, 376-387.

XPC constructs employing bulk biochemical binding assays and a fluorescence-based cellular method suggest a two-stage damage recognition model. In this model, XPC uses the BHD1/BHD2/β-turn interface to conduct a sliding search for DNA damage, followed by the more energetically costly β-hairpin insertion (264). This model is consistent with the kinetic gating mechanism of damage recognition, proposed by Min and co-workers (199).

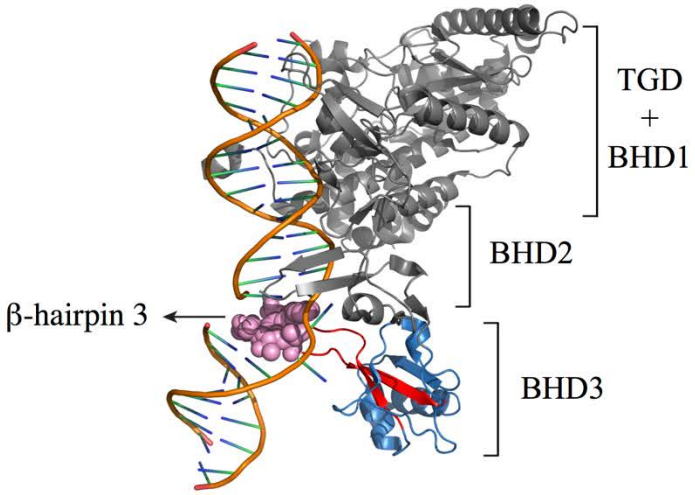

Figure 3.1: Co-crystal structure of Rad4-Rad23 with CPD-mismatch-containing DNA (PDB: 2QSG).

TGD domain and β-hairpin domains 1 and 2 or Rad4, as well as Rad23, are shown in gray, β-hairpin domains 3 in blue, with β-hairpin 3 in red. Residues 599-605 are shown in pink spheres. Adapted with permission from (279).

From the perspective of protein-DNA interactions, DNA damage recognition provides a unique example of the 'speed-stability paradox': a protein searching for target DNA sites needs to accomplish overall fast searching through a relatively smooth diffusion energy landscape while also achieving stable formation of a protein-DNA complex at the target site (deep local energy well) (218,226). We used a single-molecule DNA tightrope assay (219) and atomic force microscopy (AFM) (43,44) to (1) directly test the two-stage damage recognition model; (2) visualize how Rad4 searches for DNA damage; and (3) explore the specific role of BHD3. Using different DNA lesions and protein variants, we provide a model for how Rad4 utilizes different

structural domains to achieve damage recognition in a dynamic process. Rad4 first performs a fast initial quality check on DNA for damage detection through random diffusion and DNA bending by BHD1 and BHD2. Initial damage encounter likely triggers a protein conformational change such that it enters a more rigorous damage recognition mode characterized by constrained motion with a steeper energy landscape. This constrained motion, or 'recognition-at-a-distance,' helps to reconcile the apparent lack of specificity of Rad4 for CPDs *in vitro* (194) with its essential role of CPD repair *in vivo* (280). Finally, in a lesion and sequence specific manner, damage recognition is achieved, leading to nonmotile long-lived Rad4-DNA complexes.

3.1 RAD4-RAD23 UTILIZES A COMBINATION OF 3D AND 1D APPROACHES TO SEARCH FOR DAMAGE ON DNA

To directly visualize the Rad4-Rad23 search process, we performed single-molecule tightrope assays (43,219) using N-terminally histidine-tagged Rad4 labeled with streptavidin-conjugated quantum dots (SAQD) through biotinylated anti-histidine-tag antibody (HisAb) (Figure 3.2 and Figure B.1). The wildtype (WT) Rad4-Rad23 used here is essentially the same as that in crystal structures (Figure 3.1), spanning all four DNA-interacting domains of Rad4 (His-scRad4 101-632) and all Rad23 domains except for an internal UBA1 domain (Rad23 1-398_Δ135-299). This WT complex exhibits DNA binding behavior similar to the full-length Rad4-Rad23 complex in electrophoretic mobility shift assays (EMSAs) (88). DNA tightropes were suspended between 5 μm poly-L-lysine coated silica beads deposited randomly on a PEGylated coverslip via hydrodynamic flow using a syringe pump. SAQD-labeled Rad4-Rad23 was injected into the flow cell in the presence of DNA tightropes and observations were started immediately after flow was

stopped. For experiments performed with UV-irradiated λ-DNA containing on average one photoproduct per 2,200 bp, we detected consistent binding of Rad4-Rad23 to DNA throughout the flow cell. Over a period of ~2 hours, both motile and nonmotile complexes were observed. While some particles dissociated, we saw few arrivals during recording. This indicates that Rad4-Rad23 has a rapid on-rate through the initial 3D diffusion process, followed by 1D diffusion on DNA.

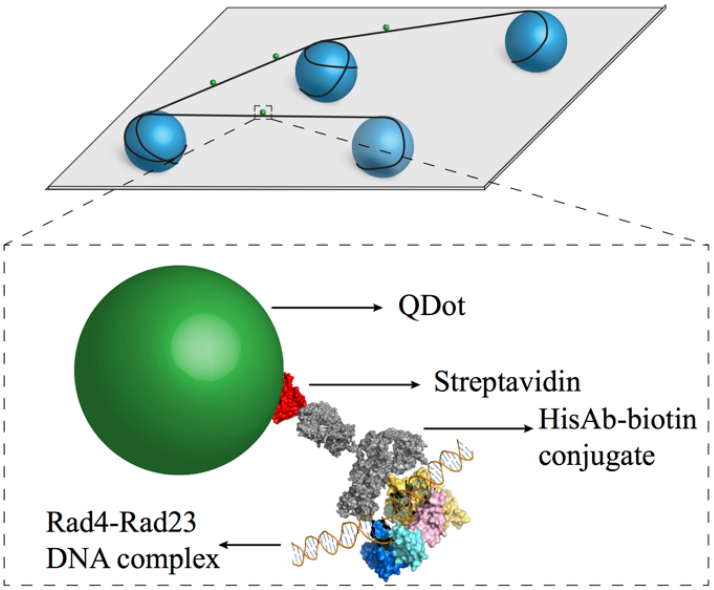

Figure 3.2: Schematics of flow cell and protein conjugation strategy.

Top: 5 μm diameter poly-L-lysine coated silica beads (blue) are deposited on polyethylene glycol treated coverslip (gray). DNA (black) is elongated and strung up across beads by flow. *Bottom*: His-tagged Rad4-Rad23 (yellow, pink, cyan, and blue) is labeled with streptavidin (red)-coated quantum dot (SAQD, green) through a His-antibody (His-Ab)-biotin conjugate (gray). Adapted with permission from (279).

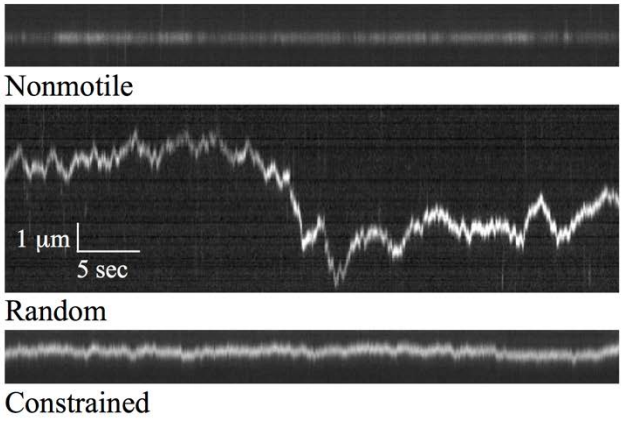

Figure 3.3: Representative kymographs.

Top: nonmotile, *middle:* random diffusion, and *bottom:* constrained motion particles. Scale bars in middle panel apply to all three kymographs. Adapted with permission from (279).

Closer examination of kymographs obtained from single particle tracking of labeled Rad4-Rad23 on DNA tightropes revealed three distinct classes of protein complex movement: nonmotile (Figure 3.3 *top*), random diffusion (Figure 3.3 *middle*), and constrained motion (Figure 3.3 *bottom*). Nonmotile particles showed no discernable movement along the trajectory of DNA (three pixels, ~500 bp, see 2.3 above) throughout the five-minute recording window. Randomly diffusing particles exhibited increasing displacement from starting positions over time; approximately 5 kbp or larger. Finally, constrained particles oscillated around certain positions on DNA and appeared restricted within ~1-2 kbp in total end-to-end displacement. On UV-irradiated λ-DNA (at 20 J/m^2), 59 ± 5% of all observed WT Rad4-Rad23 particles (n = 194) were nonmotile, 25 ± 6% diffused randomly, and 16 ± 4% underwent constrained motion (Figure 3.4).

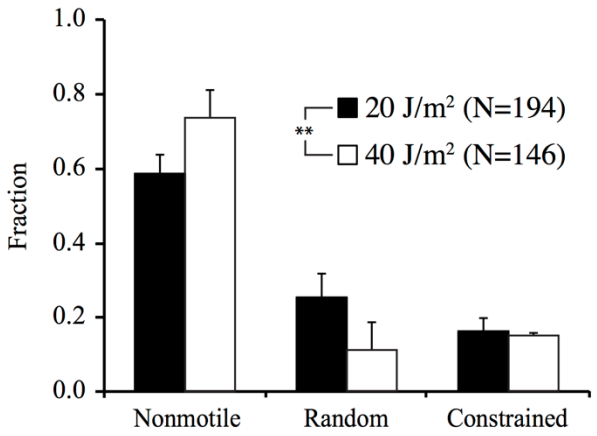

Figure 3.4: Bar graph of fractions of each observed motion type on UV-irradiated λ-DNA.

λ-DNA was irradiated with 20 J/m^2 (black bars) or 40J/m^2 (white bars) UV light (p = 0.0026, X^2 test). All bar graph data in this study are represented as weighted means ± weighted standard deviations over four to five independent experimental days. (Statistical significance *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, ****: $p \leq 0.0001$) Adapted with permission from (279).

3.2 SLIDING IS THE MAIN COMPONENT OF OBSERVED 1D DIFFUSION OF RAD4-RAD23

Among the various modes of possible protein-DNA interactions, both sliding and hopping along DNA are perceived as 1D diffusion. However, their mechanisms are fundamentally different. Proteins undergoing 1D sliding maintain contacts with DNA continuously through a corkscrew motion along the helical path. In contrast, hopping involves microscopic dissociation from and rebinding to the same piece of DNA, without macro-dissociating into solution (225,281). During periods of microscopic dissociation, proteins are removed just far enough from DNA such that cations can re-condensate onto the phosphate backbone of DNA (282). Since stronger electrostatic screening negatively affects affinities of nonspecific binding events, diffusion by sliding should be relatively insensitive to changes in salt concentrations, while hopping particles are expected to

show an increase in their diffusion coefficients (*D*) as the distances between hops grow under higher salt conditions (222,247). Thus to differentiate between the two mechanisms, we repeated experiments with WT proteins on 20 J/m^2 UV-irradiated λ-DNA at 100 mM and 150 mM NaCl (n = 172 and 169, respectively), compared with 75 mM NaCl. Across three salt conditions, nonmotile fractions remained similar (~60%, Figure 3.5). Additionally, distributions of $log_{10}D$ (see 2.3) of over 90% of all motile particles, random and constrained, were normal with similar means and standard deviations in all three salt conditions (Figure 3.6). The apparent insensitivity of diffusion coefficients to higher ionic strengths therefore indicates that under these salt concentrations (75-150 mM NaCl), the majority (>90%) of all motile Rad4-Rad23 slide as they 1D diffuse along DNA. Interestingly, we also observed that at 150 mM NaCl, the number of particles possessing the highest diffusion coefficients (D ~ 0.1 μm^2/s) increased by ~10% of the total (n = 65, Figure 3.6 *bottom*) compared to results seen in 75 and 100 mM NaCl.

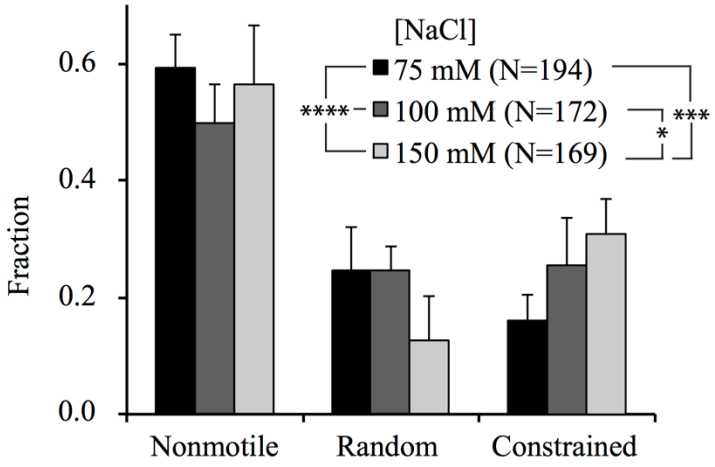

Figure 3.5: Distributions of observed motion types at different salt concentrations.

Data at 75 mM NaCl reproduced from **Figure 3.4** (p = 0.0005, X^2 test). Adapted with permission from (279).

To investigate the diffusion nature in more detail, we analyzed the anomalous diffusion exponent α for all motile particles (see 2.3). α is expected to be ~ 1 for randomly diffusing particles

and < 1 for constrained particles. These analyses show that the fraction of particles undergoing constrained motion increased with more physiological salt concentrations (Figure 3.5). Increasing ionic strength to 150 mM NaCl resulted in an increased population that exhibited $\alpha < 1$ (Figure 3.6 *bottom*). Finally, comparing relationships between anomalous diffusion exponent α and diffusion coefficient D across three salt concentrations (Figure 3.6), the fast diffusers ($D \sim 0.1$ μm^2/s) seen at 150 mM NaCl appear to have arisen from the random diffusion population (Figure 3.6 *bottom*). Taken together, these data support the hypothesis that particles undergoing constrained motion are indeed sliding on DNA; their behavior remained relatively unchanged when challenged with higher salt. On the other hand, we cannot exclude the possibility that some randomly diffusing particles may undergo hopping at higher salt concentrations.

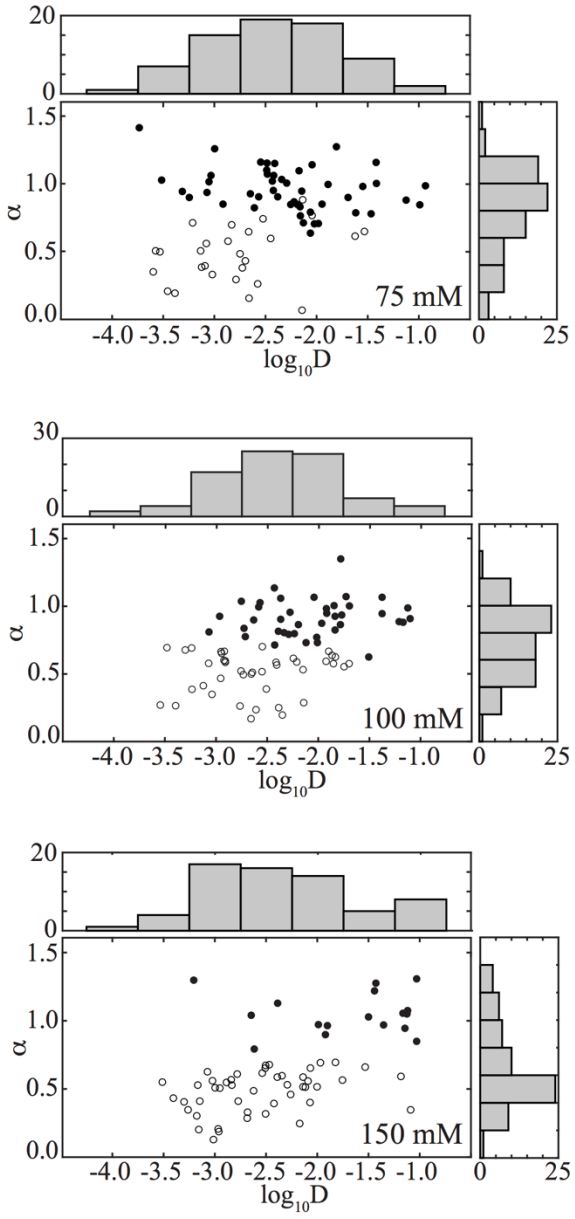

Figure 3.6: Anomalous diffusion exponent (α) vs. diffusion coefficient ($log_{10}D$) of Rad4-Rad23 diffusing at different salt concentrations.

Random (filled circles) and constrained (empty circles) particles at 75 mM (*top*), 100 mM (*middle*), and 150 mM (*bottom*) NaCl. Distributions of diffusion coefficients $log_{10}D$ and anomalous diffusion exponents α are plotted above and to the right of each scatter plot, respectively. Adapted with permission from (279).

3.3 RAD4-RAD23 EXHIBITS LESION-SPECIFIC DAMAGE RECOGNITION

It is interesting to note that ~60% of WT Rad4 observed on UV-irradiated λ-DNA were nonmotile (Figure 3.4, black bars). Increasing the UV dose two-fold increased the percentage of nonmotile particles and decreased random movers (Figure 3.4, white bars). UV irradiation induces a mixture of CPDs and 6-4PPs at about a 3:1 ratio (283). Rad4-Rad23 binds poorly to CPDs *in vitro*, while possessing an order of magnitude higher affinity toward the helix-distorting 6-4PPs (194). We thus characterized the Rad4-Rad23 behavior on DNA substrates that harbor one type of DNA lesion in the same repeating sequence context. To this end, we made long DNA-damage arrays by tandemly ligating multiple linearized plasmids, each contained either one CPD or one fluorescein-modified deoxythymidine (Fl-dT) per 2,030 bp, as previously described (43). Rad4 binds tightly to Fl-dT (196), making it a model substrate with high specificity (Figure B.2 *top*). As expected, SAQD-labeled Rad4-Rad23 formed arrays of nonmotile complexes when introduced into flow cells in the presence of Fl-dT DNA tightropes (Figure 3.7), with inter-particle spacing being integer-multiples of 2 kbp (Figure B.3). Overall, 80 ± 18% of all Rad4-Rad23 particles on Fl-dT DNA were nonmotile, while random and constrained movers represented 12 ± 11% and 8 ± 12%, respectively (Figure 3.9, n = 211). In comparison, the nonmotile Rad4-Rad23 population was reduced by 2-fold to 42 ± 10% on CPD-containing DNA damage arrays (n = 106). The fraction of random movers on CPD substrates remained at 14 ± 6%, whereas that of constrained particles increased more than five-fold to 44 ± 4% compared to Fl-dT (Figure 3.8). In contrast, Rad4-Rad23 behavior on undamaged DNA was statistically different from that on CPD arrays (Figure 3.9, $p = 0.0187$, X^2 test) and important differences were noted in motion types: more particles (26 ± 12%) diffused randomly and less (27 ± 8%) underwent constrained motion on undamaged DNA as compared to CPDs (Figure 3.9). These results suggest that constrained motion is directly due to recognition of

CPDs. Furthermore, only 6% of all particles on DNA bound at two positions on undamaged DNA tightropes; over 55% and 41% of particles bound to Fl-dT- and CPD-containing substrates at regular intervals consistent with inter-lesion distance of 2 kbp (Figure B.3). The presence of Rad4-Rad23 arrays on CPD substrates and the lack thereof on undamaged DNA argue that, in our tightrope setup, the protein senses the relatively minor helical distortion caused by this lesion.

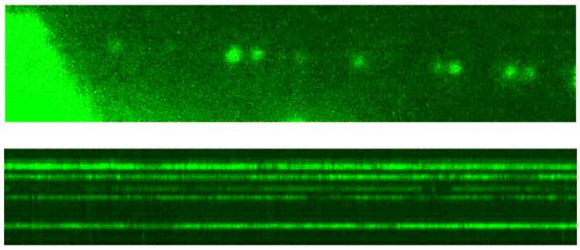

Figure 3.7: Single frame (top) and kymograph (bottom) of quantum dot-labeled Rad4-Rad23 particles assembled in an array on Fl-dT-containing DNA.

Adapted with permission from (279).

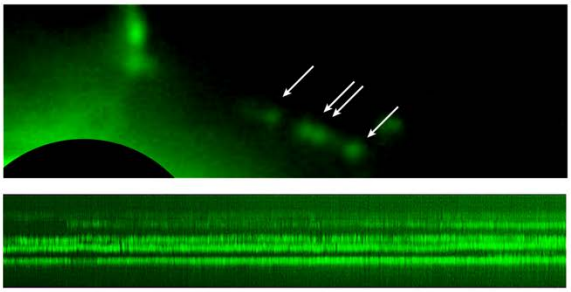

Figure 3.8: Single frame (top) and kymograph (bottom) of quantum dot-labeled Rad4-Rad23 particles assembled in an array on CPD-containing DNA.

Rad4-Rad23 particles are indicated by white arrows. A black semi-circle was added in the lower left corner to mask out the bright bead. Adapted with permission from (279).

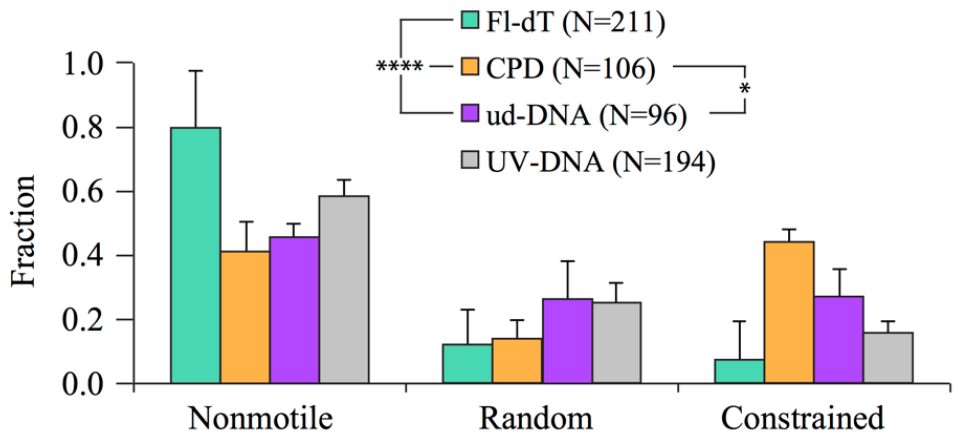

Figure 3.9: Distributions of motion types of WT Rad4-Rad23 observed on DNA damage arrays.

Lesion-dependent behavior of WT Rad4-Rad23 on DNA damage arrays (Fl-dT – green, CPD – orange, undamaged DNA – purple, 20 J/m^2 UV-irradiated λ-DNA – gray). Adapted with permission from (279).

3.4 TRUNCATIONS IN THE β-HAIRPIN DOMAIN 3 (BHD3) OF RAD4 INCREASE CONSTRAINED MOTION

Co-crystal structure of Rad4-Rad23 bound to model DNA lesions show that the tip of β-hairpin 3 of Rad4 inserts into the DNA duplex at the lesion site, suggesting that this structural component is important for DNA damage recognition (Figure 3.1). Previous EMSAs have also demonstrated that deletions of the β-hairpin tip (Δβ-hairpin3, His-scRad4 101-632_Δ599-606) or β-hairpin domain 3 (ΔBHD3, His-scRad4 101-540) abolished lesion-specific binding of Rad4-Rad23 (88,199). We further investigated the role of the β-hairpin 3 motif in Rad4's lesion recognition by examining the diffusion behavior of Δβ-hairpin3 (β-hairpin 3 tip deletion, His-scRad4 101-632_Δ599-606) and ΔBHD3 (β-hairpin 3 domain deletion, His-scRad4 101-540) (Figure 3.1) on UV-irradiated λ-DNA. Both mutants showed a decrease in nonmotile particles (30-40%) compared to WT (~60%, Figure 3.10). An overall upward trend in the random diffusion population also

corresponded to the increasing loss of residues within BHD3 in these two mutants (25% for WT, 33% for Δβ-hairpin3, and 39% for ΔBHD3; Figure 3.10). The population shift towards faster random diffusion for the ΔBHD3 mutant is also reflected in plots of anomalous diffusion exponents versus diffusion coefficients (Figure 3.11). Interestingly, the fraction of Δβ-hairpin3 undergoing constrained motion is almost double of WT or ΔBHD3.

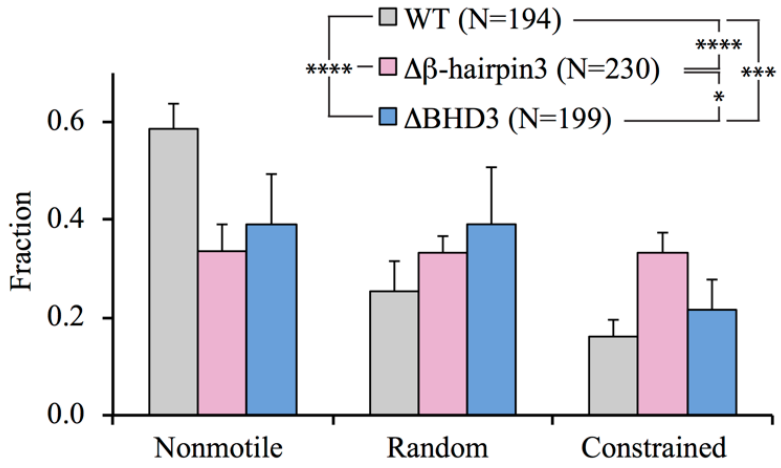

Figure 3.10: Distributions of observed motion types from Rad4 WT and mutants.
Adapted with permission from (279).

X^2 analysis of WT and both mutants shows that distributions of motion types were indeed affected by deletions of the damage-sensing β-hairpin 3 ($p < 0.0001$, Figure 3.10). Histograms of diffusion coefficients and anomalous diffusion exponents from all motile particles also show that while the Δβ-hairpin3 diffusion coefficient was similar to WT, deletion of the full domain (ΔBHD3) caused ~25% of complexes to diffuse significantly faster (Figure B.4). Finally, as compared to WT, both Δβ-hairpin3 and ΔBHD3 mutants appear to be more prone to dissociation (15%, 24%, 31%, respectively, Figure 3.12). However, at low particle counts, neither dissociation kinetics nor mean lifetimes of the dissociating proteins were significantly different across three protein variants,

shown by Mantel-Cox log-rank test of survival curves (Figure 3.13, $p > 0.5$) and one-way ANOVA of off-rates obtained from single-exponential fitting of lifetime histograms (Figure 0.5, $p > 0.2$).

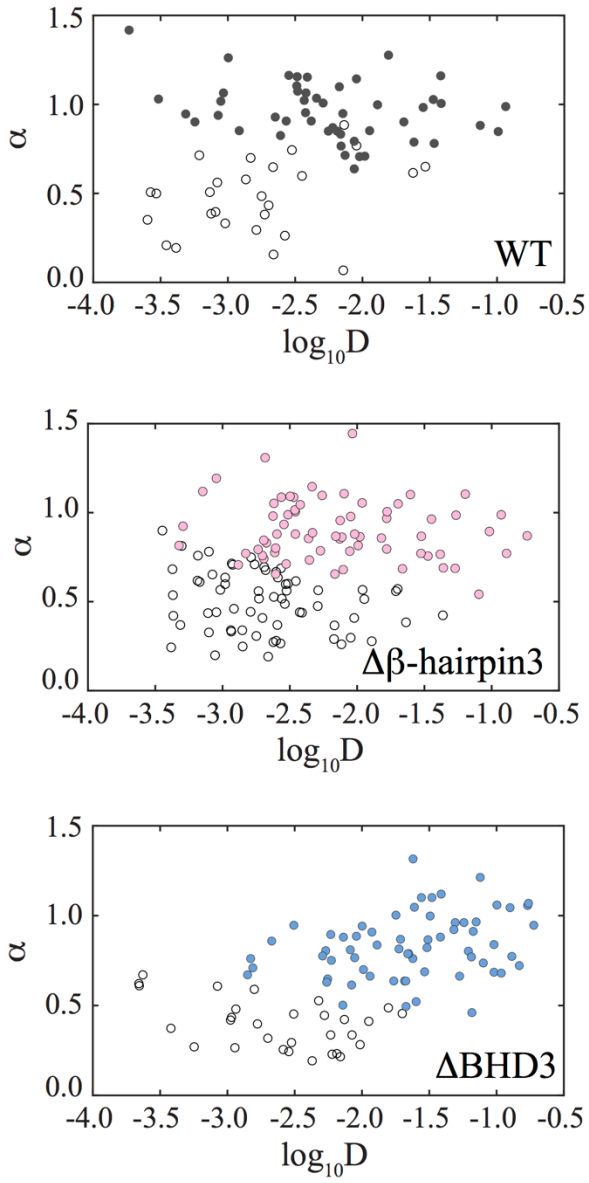

Figure 3.11: Anomalous diffusion exponent (α) vs. diffusion coefficient (log10D) of Rad4-Rad23 WT and mutants diffusing on UV-irradiated λ-DNA.

Random (filled circles) and constrained (empty circles) particles of WT (*top*), Δβ-hairpin3 (*middle*), and ΔBHD3 (*bottom*) on 20 J/m^2 UV-irradiated λ-DNA. Adapted with permission from (279).

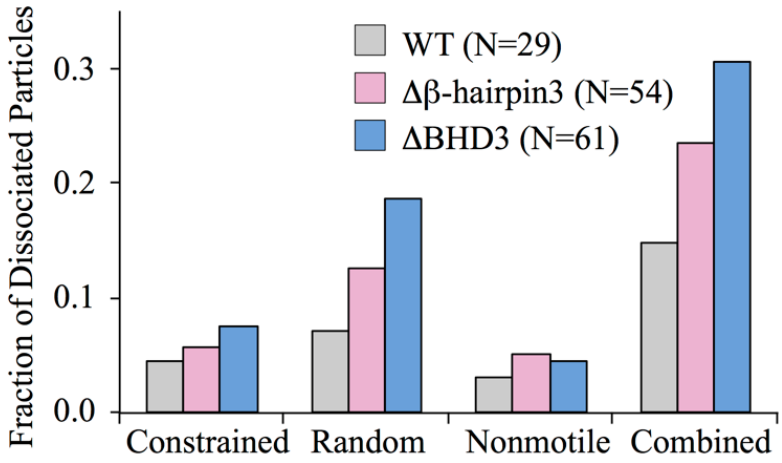

Figure 3.12: Dissociating particles as fractions of total particles observed increase with larger deletions in Rad4 BHD3 sequence.

Adapted with permission from (279).

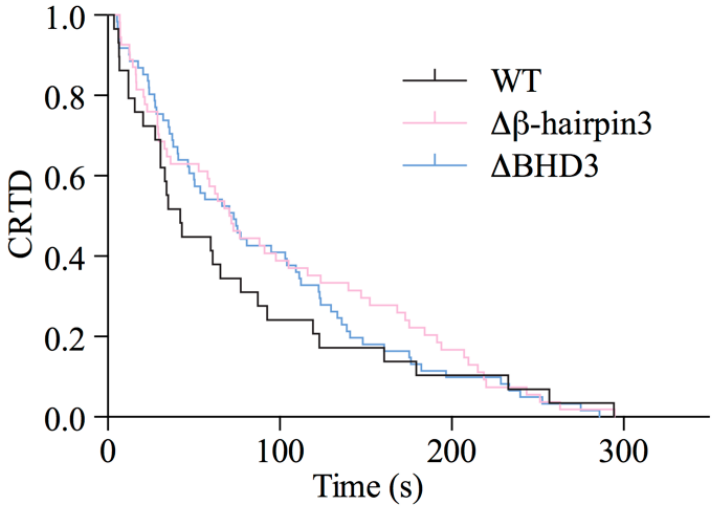

Figure 3.13: Cumulative residence time distribution (CRTD) plot of lifetimes of Rad4 WT and mutants that dissociated during observation.

Adapted with permission from (279).

Further analysis of these protein variants on DNA damage arrays revealed distinct behavior of the deletion mutants compared to WT on Fl-dT substrates (Figure 3.14 *top*, $p < 0.01$, X^2 test). While Δβ-hairpin3 behaved similarly to WT on CPD substrates, ΔBHD3 showed significant increase in constrained motion (Figure 3.14 *bottom*, $p < 0.05$, X^2 test).

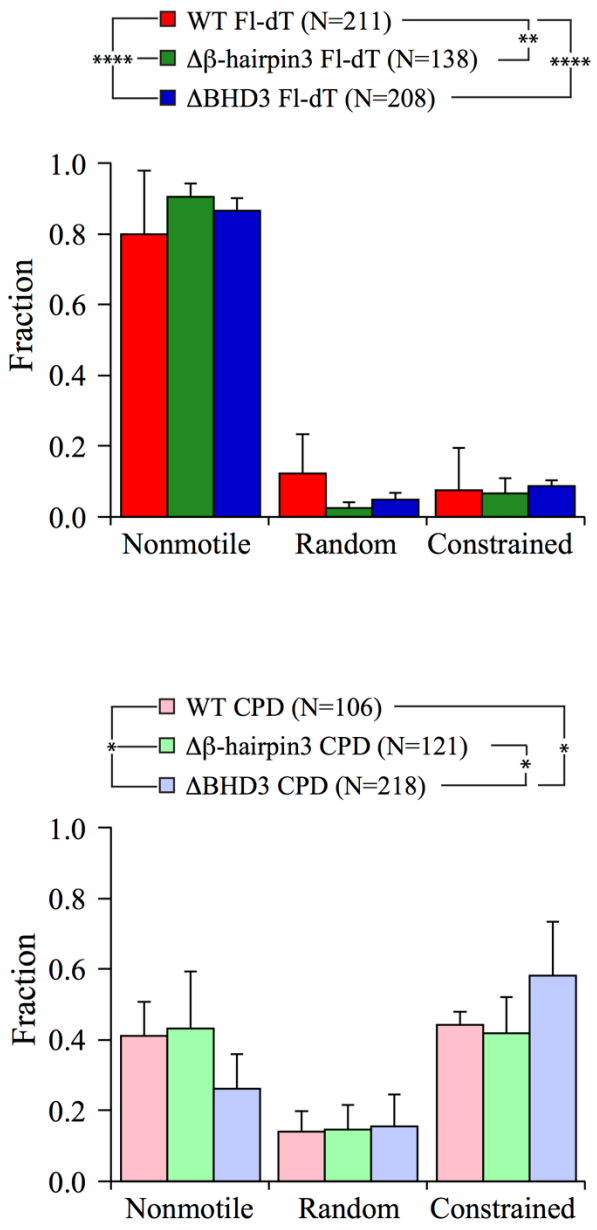

Figure 3.14: Distributions of motion types of WT, Δβ-hairpin3, and ΔBHD3 observed on DNA damage arrays.

Top: Fl-dT-containing substrates (WT – red, Δβ-hairpin3 – green, ΔBHD3 – blue). *Bottom*: CPD-containing substrates (WT – pink, Δβ-hairpin3 – mint, ΔBHD3 – lavender). WT data reproduced from Figure 3.9. Adapted with permission from (279).

3.5 RAD4 VARIANT LACKING β-HAIRPIN DOMAIN 3 (ΔBHD3) IS CAPABLE OF SPECIFIC BINDING AND DNA BENDING TO FL-dT-CONTAINING DNA FRAGMENTS

The co-crystal of DNA-bound Rad4-Rad23 indicates that DNA binding by WT causes a kink in the DNA of about 42° (Figure 3.1) (88). Having shown that the ΔBHD3 variant forms stable complexes on UV-irradiated λ-DNA tightropes and binds to Fl-dT-containing a short DNA fragment in fluorescence anisotropy experiments (Figure B.2 and Table B.1), we asked if DNA bending is also a feature of binding by this mutant. Using AFM we studied the bending of a 538-bp dsDNA fragment that contains a Fl-dT lesion at 30% of the contour length from one end (Figure 3.15). Naked DNA was bent by only 4 ± 32° (n = 245, Figure B.6), whereas we observed that WT specifically bound at 32 ± 13% contour length (n = 335, Figure 3.16 *top*) bent the DNA 43 ± 24° (n = 189, Figure 3.16 *bottom*); remarkably similar to specifically bound ΔBHD3 (31 ± 10% contour length, n = 148, Figure 3.17 *top*) which bent DNA 36.5 ± 29.1° (n = 101, Figure 3.17 *bottom*). Lastly, consistent with our previous report on WT protein binding to undamaged DNA fragments (199), both WT and ΔBHD3 induced bending in DNA even when not specifically bound to the Fl-dT lesion (white bars, Figure 3.16 *bottom* and Figure 3.17 *bottom*).

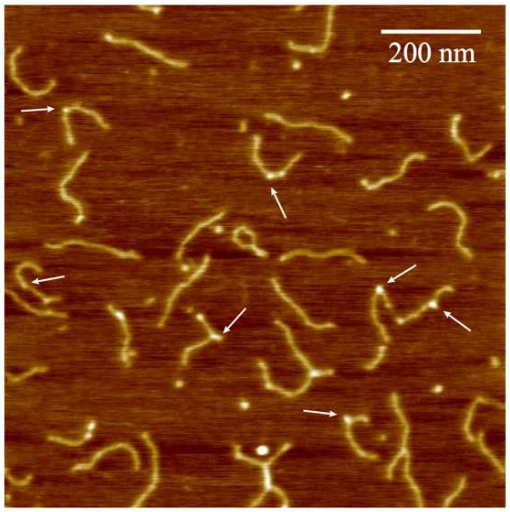

Figure 3.15: Representative AFM image of ΔBHD3 bound to Fl-dT-containing DNA fragments.

White arrows highlight representative binding events scored in data analysis. Adapted with permission from (279).

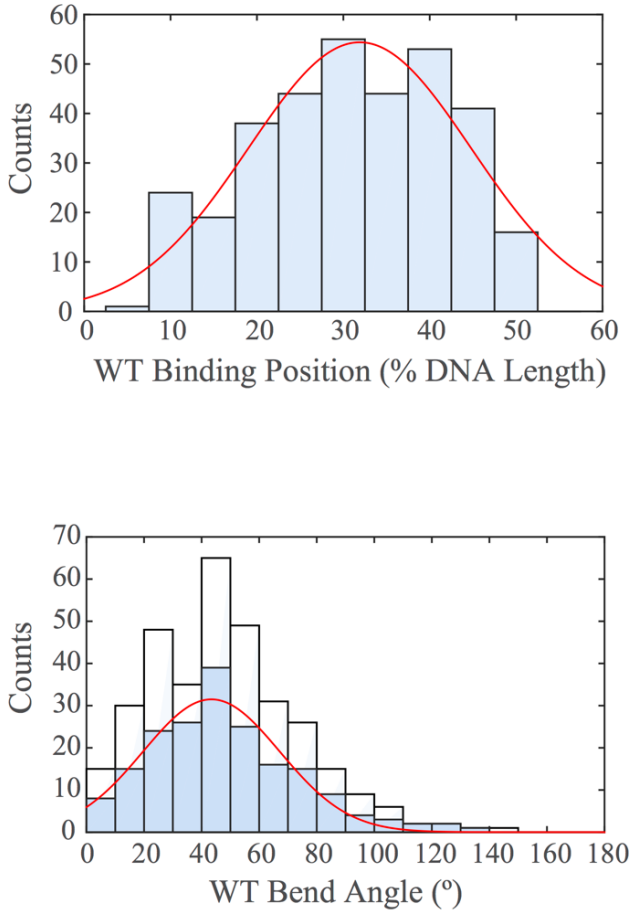

Figure 3.16: Distributions of WT binding positions and bend angles.

Top: Histogram and Gaussian fitting (red curve) of WT binding positions (32 ± 13%, n = 335) on DNA fragment in terms of percentage of total contour length measured from one end.
Bottom: Histogram of DNA bend angles at all internal WT binding sites (white, n = 335). Histogram (blue) and Gaussian fitting (red curve) of DNA bend angles (43 ± 24°, n = 189) at WT proteins specifically bound between 20% and 40%.
Adapted with permission from (279).

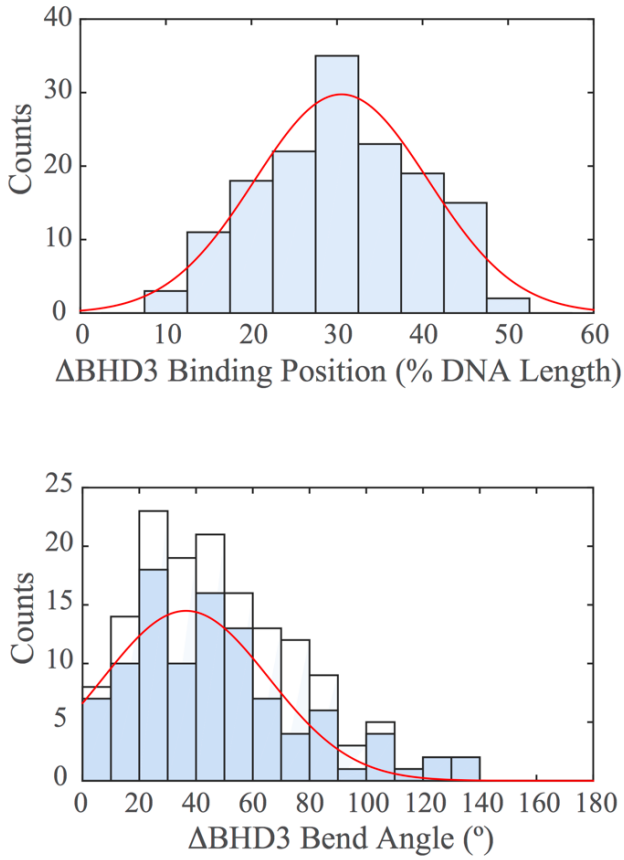

Figure 3.17: Distributions of ΔBHD3 binding positions and bend angles.

Top: Histogram and Gaussian fitting (red curve) of ΔBHD3 binding positions (31 ± 10%, n = 148) on DNA fragment in terms of percentage of total contour length measured from one end.
Bottom: Histogram of DNA bend angles at all internal ΔBHD3 binding sites (white). Histogram (blue) and Gaussian fitting (red curve) of DNA bend angles (37 ± 29°, n = 101) at ΔBHD3 specifically bound between 20% and 40%.
Adapted with permission from (279).

3.6 DELETIONS OF C-TERMINAL REGIONS IN RAD4 CONFER VARYING DEGREES OF UV SENSITIVITY AND REPAIR IN *S. CEREVISIAE*

Since the seven amino acid deletion in β-hairpin 3 resulted in more constrained motion on UV-irradiated DNA, we next tested whether this variant promotes efficient repair *in vivo*. We thus conducted UV survival and DNA repair assays on yeast strains carrying different Rad4 mutants

with deletions and truncations around β-hairpin 3 within the *RAD4* locus (Figure 3.1 and Figure 3.18). Notably, deletion of the seven amino acid tip of β-hairpin 3 (*rad4* Δ599-605) showed WT levels of UV resistance (Figure 3.19, compare pink and black lines). Removal of the entire β-hairpin 3 from stem to tip (Figure 3.19, *rad4* Δ590-615, red) led to increased UV sensitivity comparable to deleting the entire β-hairpin domain 3 (Figure 3.19, *rad4* Δ541-632, blue). However, both of these constructs were considerably more UV resistant than the *rad4*Δ strain. Finally, cells with truncation from BHD3 to the C-terminus (Figure 3.19, *rad4* Δ541-Cterm, green) are as UV sensitive as the *rad4*Δ strain (Figure 3.19, *rad4*Δ, orange). Protein expression levels of FLAG-tagged Rad4 WT and mutants were probed with α-FLAG antibody (Figure 3.20). Western blotting data suggest that removal of the entire β-hairpin 3 destabilized the protein and could contribute to some of the observed UV sensitivity due to reduced Rad4 protein levels.

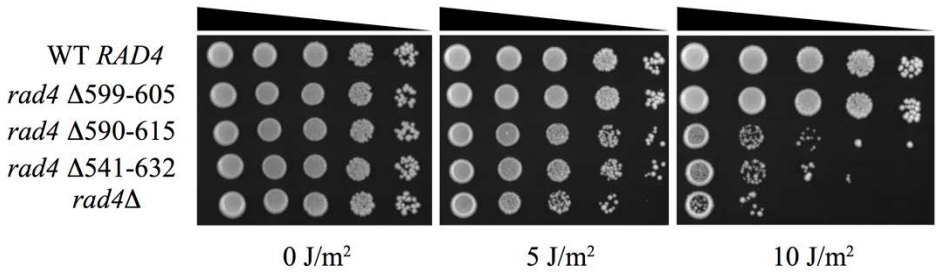

Figure 3.18: Serial dilutions of yeast cells (BY4742) expressing different 3xFLAG-tagged Rad4 variants on YPD plates, 72 hours after UV irradiation.

Adapted with permission from (279).

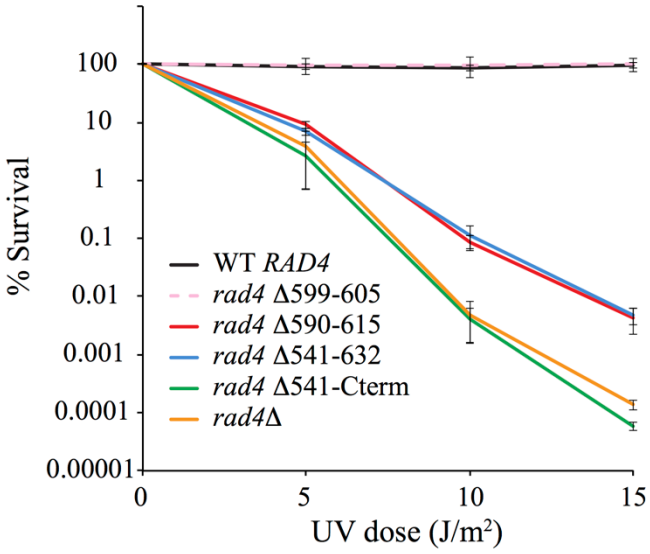

Figure 3.19: Quantitative UV-survival of yeast cells (BY4741) expressing different untagged Rad4 variants.

WT *RAD4* – black, *rad4* Δ599-605 (Δβ-hairpin3) – pink dashed, *rad4* Δ590-615 – red, *rad4* Δ541-632 (ΔBHD3) – blue, *rad4* Δ541-cterm – green, *rad4*Δ – orange. Adapted with permission from (279).

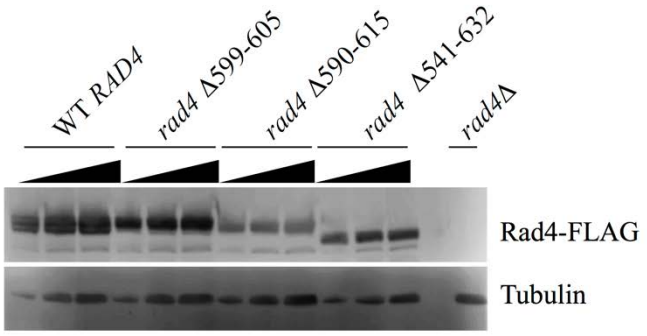

Figure 3.20: Expression levels of 3xFLAG-tagged Rad4 variants detected with anti-FLAG antibody. Adapted with permission from (279).

To investigate whether the mutant lacking the seven amino acid tip of β-hairpin 3 (Δβ-hairpin3, *rad4* Δ599-605), which undergoes more constrained motion on UV-irradiated λ-DNA, also showed WT levels of photoproduct repair, we used two different experimental approaches: 1) T4 phage pyrimidine dimer glycosylase (Endo V) incisions on genomic DNA (Figure 3.21 and

Figure 3.22), and 2) antibody slot blots of total genomic DNA for CPD and 6-4PP (Figure B.7). These data indicate that loss of seven amino acids from the tip of β-hairpin 3 did not affect the rates of CPD or 6-4PP repair. However, consistent with UV survival data, larger deletions of β-hairpin domain 3 resulted in loss of photoproduct removal.

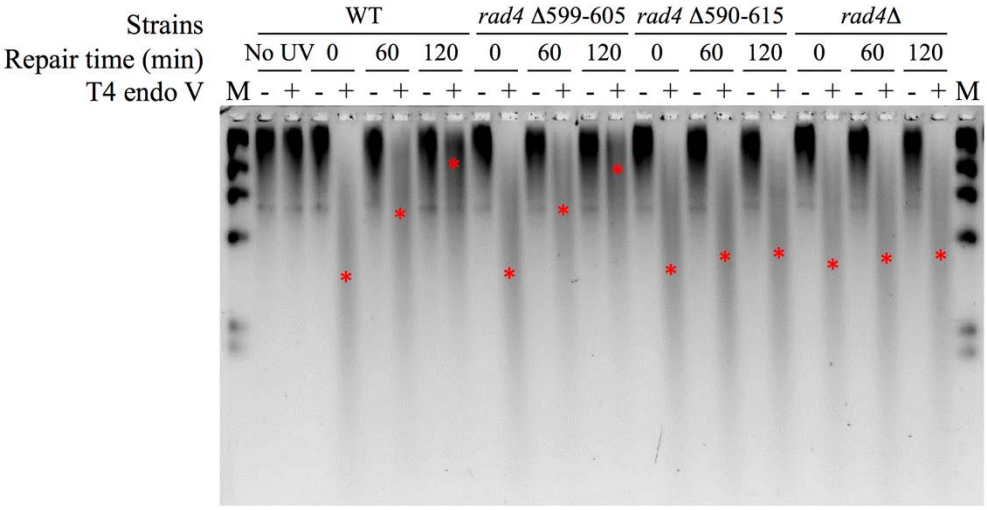

Figure 3.21: Genomic DNA of yeast cells after UV irradiation and recovery digested with T4 endo V, separated on alkaline agarose gel, and detected with SYBR Gold.

Approximate positions of the ensemble average size of DNA in each lane are denoted with red asterisks (*). DNA marker (M, λ-DNA-HindIII) was loaded in the left- and right-most lanes. Adapted with permission from (279).

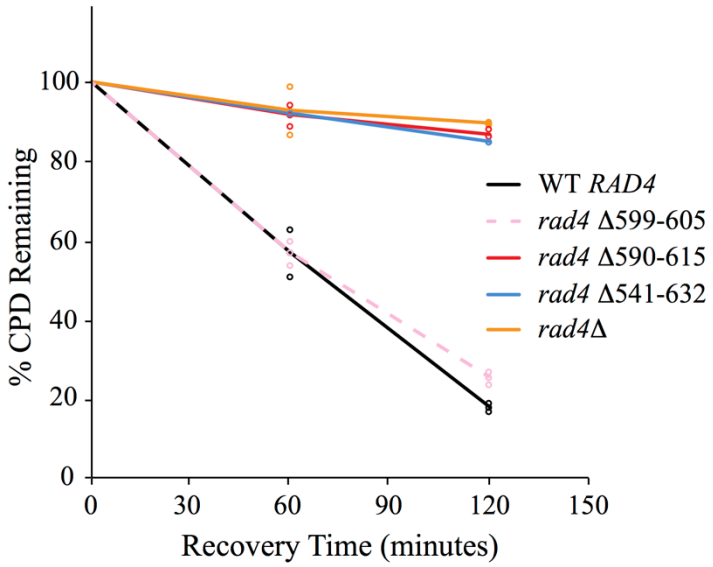

Figure 3.22: Quantitative rates of CPD removal of yeast cells (BY4741) expressing different untagged Rad4 variants, determined by T4 endo V digestion.

Color scheme same as in Figure 3.19. Adapted with permission from (279).

3.7 DISCUSSION

Here, we employed single-molecule methods to uncover the dynamic nature of the Rad4-Rad23 damage recognition process. We found that Rad4-Rad23 forms stable protein-DNA complexes or slides on DNA one-dimensionally to search for damage. In addition to random 1D diffusion, we showed that some Rad4-Rad23 molecules exhibited constrained motion (~1-2 kbp) around damage sites in a lesion-dependent manner. This apparent subdiffusive behavior was also influenced by deletions made in the β-hairpin domain 3 of Rad4. Surprisingly, AFM experiments revealed that Rad4-Rad23 lacking β-hairpin domain 3 (ΔBHD3) binds specifically to Fl-dT, while inducing a bend in DNA similar to WT binding, suggesting that BHD3 is not directly involved in initial damage detection or DNA bending at Fl-dT modified sites. Furthermore, we demonstrated that

Rad4 mutant Δβ-hairpin3 lacking seven amino acids (FERGSTV) at the tip of the β-hairpin 3 caused a 2-fold increase in constrained motion on DNA tightropes, while maintaining WT levels of UV resistance as well as CPD and 6-4PP removal in yeast. This work examining long range motions of Rad4/Rad23 on DNA strongly supports a model in which Rad4 uses constrained motion around CPD sites as this "recognition-at-a-distance" mechanism allows efficient repair.

3.7.1 Alternative Damage Recognition Mechanism for Sub-Optimal Substrates through Constrained Motion by Rad4-Rad23

Our working model for Rad4-Rad23 damage recognition suggests that constrained motion represents an intermediate interrogation step. UV-irradiation induces structurally distinct lesions that are differentially bound by Rad4-Rad23 (194). CPDs are less distorting to the DNA helix than 6-4PPs (89) and Fl-dT likely intercalates between DNA base pairs (284). Robust stable and specific binding of Rad4-Rad23 to Fl-dT-containing DNA damage arrays (Figure 3.9) is consistent with our fluorescence anisotropy measurements (Figure B.2 and Table B.1), as well as previous reports (196). In contrast, the subdiffusive population of WT Rad4-Rad23 on CPD-containing damage arrays increased 5-fold compared to Fl-dT at the expense of stably bound particles (Figure 3.9). This observation substantiates the idea that in addition to stable binding at specific lesions; Rad4-Rad23 can effectively convey damage recognition through constrained motion around damage sites, particularly at the 'sub-optimal' weakly distorting CPDs.

Observation of distinct constrained motion of Rad4-Rad23 around CPD sites may also help reconcile the discrepancy in reported roles of Rad4 in CPD removal, where Rad4 is incapable of recognizing CPDs *in vitro* (194), yet indispensable for removal of thymine dimers *in vivo* (280). Past biochemical studies characterizing binding of Rad4-Rad23 utilized short (~100 bp) damage-

containing DNA fragments (194,196). Since the protein exhibits oscillatory motion of ~1-2 kbp around CPD sites on naked DNA, one plausible explanation for Rad4-Rad23's apparent lack of specificity towards CPD is that Rad4-Rad23 may dissociate from ends of such short DNA fragments *in vitro*. Our data from experiments on both UV survival as well as CPD and 6-4PP repair kinetics support a model in which Rad4 recruits downstream repair factors while undergoing constrained motion on genomic DNA. Since eukaryotic DNA is organized into chromatin with one nucleosome every 147 bp (285), and only 1-2 nucleosomes are removed during NER (286), this constrained motion would be of significantly shorter ranges than measured on DNA tightropes, making site-specific recruitment of downstream proteins even more efficient.

3.7.2 BHD3-Independent DNA Bending as an Initial Quality Check by Rad4

Our data on DNA bending by specifically bound WT Rad4-Rad23 (Figure 3.16 *bottom*, blue bars) are consistent with the crystal structure (88), as well as a previous study on XPC binding to cholesterol damage (277). We also observed bending in DNA produced by non-specifically bound WT proteins (Figure 3.16 *bottom*, white bars), consistent with our previous report (199). β-hairpin 3, seen inserted between DNA strands at site of lesion in the crystal structure (Figure 3.1), has been hypothesized to be crucial for damage recognition (88). Remarkably, we have shown that Rad4 lacking the entire β-hairpin domain 3 (ΔBHD3) was capable of specific binding to the Fl-dT lesion (Figure 3.17 *top*) and induced a bend in DNA comparable to that caused by WT protein both at damage sites (Figure 3.17 *bottom*, blue bars) and on non-specific undamaged sequences (Figure 3.17 *bottom*, white bars). These data suggest that Rad4 checks the integrity of DNA using β-hairpin domains 1 and 2. This initial recognition is thus independent of the energetically costly insertion of β-hairpin 3 into DNA. Since β-hairpin insertion is slow and rate-limiting when binding

to mismatch DNA (199), DNA bending may serve as a rapid initial quality check on a much faster timescale, e.g. during linear diffusion on DNA where the residence time at each base pair is ~10 μs (see 2.11). At strongly helix-distorting lesions such as a Fl-dT or 6-4PP, bending/twisting of DNA could lead to spontaneous base-flipping (287) and trapping of Rad4-Rad23 in an energy minimum such that robust protein-DNA complexes are formed (199,200). In contrast, at damage sites with minimal helical distortions, where β-hairpin 3 insertion and subsequent protein-DNA complex stabilization are less attainable, TFIIH and Rad14 (XPA) may be relied on more heavily, as in the recently proposed 'tripartite DNA lesion recognition and verification' process (137). Increased involvement of downstream NER factors would also help explain the slower repair rate of CPDs (288). Finally, in addition to damage recognition by Rad4-Rad23, the Rad7-Rad16 complex, known to be essential for dimer removal in silenced genes and again contribute to ~20-30% of CPD repair in the non-transcribed strands of active genes (280), have also been implicated in some cases to function as a putative damage sensor (289,290).

3.7.3 A Dynamic Multi-Step Damage Recognition Model

Subdiffusion of macromolecules in biological systems has been observed previously (43,261,266,291-293). Rad4 showed increased constrained motion at physiological salt concentrations (Figure 3.5 and Figure 3.6), which could be due to the favorable hydrophobic interactions between aromatic side-chains (F556, F597, and F599) and DNA bases at elevated ionic strengths. A recent molecular dynamics simulations study has identified that F556, F597, and F599 form a Phe 'flipping path' in BHD3, facilitating β-hairpin 3 insertion (201). Because the correct orientation of F599 during base flipping was essential in allowing complete insertion of the hairpin, loss of F599 in Δβ-hairpin3 may impede or abort the insertion process, resulting in

increased constrained motion rather than formation of stable complexes. The hypothesis that Rad4-Rad23 undergoing constrained motion remains 'repair competent' is corroborated by the finding that yeast carrying the protein variant lacking seven amino acids at the tip of β-hairpin 3 (Δβ-hairpin3) are as UV-resistant as WT (Figure 3.18) and show the same rates of CPD and 6-4PP removal (Figure 3.22 and Figure B.7). Any decrease in recognition and repair due to reduced levels of stable binding of Δβ-hairpin3 is apparently compensated by this 'recognition-at-a-distance' achieved through subdiffusion of the protein around the lesion. We thus envision Rad4-Rad23 as a first responder to arrive at the scene of an accident, able to direct other emergency workers to the site without being directly on the scene. Indeed, 'recognition-at-a-distance' may be applicable to a wide range of proteins that need to achieve target binding and signal for downstream processes, such as those involved in replication and transcription, during which 'molecular traffic jams' could occur (294). This mechanism would allow weakly interacting or sub-optimal target sites to be recognized and acted upon, while reducing potential steric hindrance or target site occlusion problems between the tightly bound recognition proteins and subsequent factors that need access to the targets. Mismatch repair proteins that dissociate from mismatched bases in an ATP-dependent manner to recruit the next proteins may also fall into this general category (261).

The balance in maintaining speed and specificity to target search and recognition by DNA binding proteins has been subject to both theoretical and experimental studies (218). In a previously established two-state model, a protein is considered to have two conformations, one that allows rapid diffusion on a smooth energy landscape, and the other that binds to target with a rugged landscape required for high specificity (226). Similar to the previously proposed conformational proofreading mechanism (43,295), our data on diffusion of Rad4 WT and mutants support a dynamic model with multiple intermediate states that utilize different structural domains

of the protein to achieve efficient damage recognition (Figure 3.23). We calculate that the energy barrier to free diffusion is $\sim 1.60\ k_B T$ and $\sim 0.37\ k_B T$ for WT undergoing constrained motion and random diffusion, respectively (see 2.10). DNA bending and other interactions between β-hairpin domain 2 and DNA likely contribute to the ruggedness of energy landscape during subdiffusion. Base flipping and stabilization of flipped-out bases following β-hairpin 3 insertion, which amounts to $\sim 5.7\ k_B T$ (199,201), make further contributions towards and exceeding the theoretical minimum energy difference requirement at target sites ($\sim 5.7\ k_B T$ for yeast genome, see 2.12) (226,273). Overall, assuming WT diffusional behavior is observed in a yeast cell nucleus containing $\sim 1.2 \times 10^7 bp$ of genomic DNA and ~870 copies of Rad4, we can estimate the shortest possible time needed to search the genome from the typical range of motion of a Rad4-Rad23 molecule and its measured average lifetime. Such calculation yields a lower limit of genome search time of roughly 2 – 3 minutes (see 2.13).

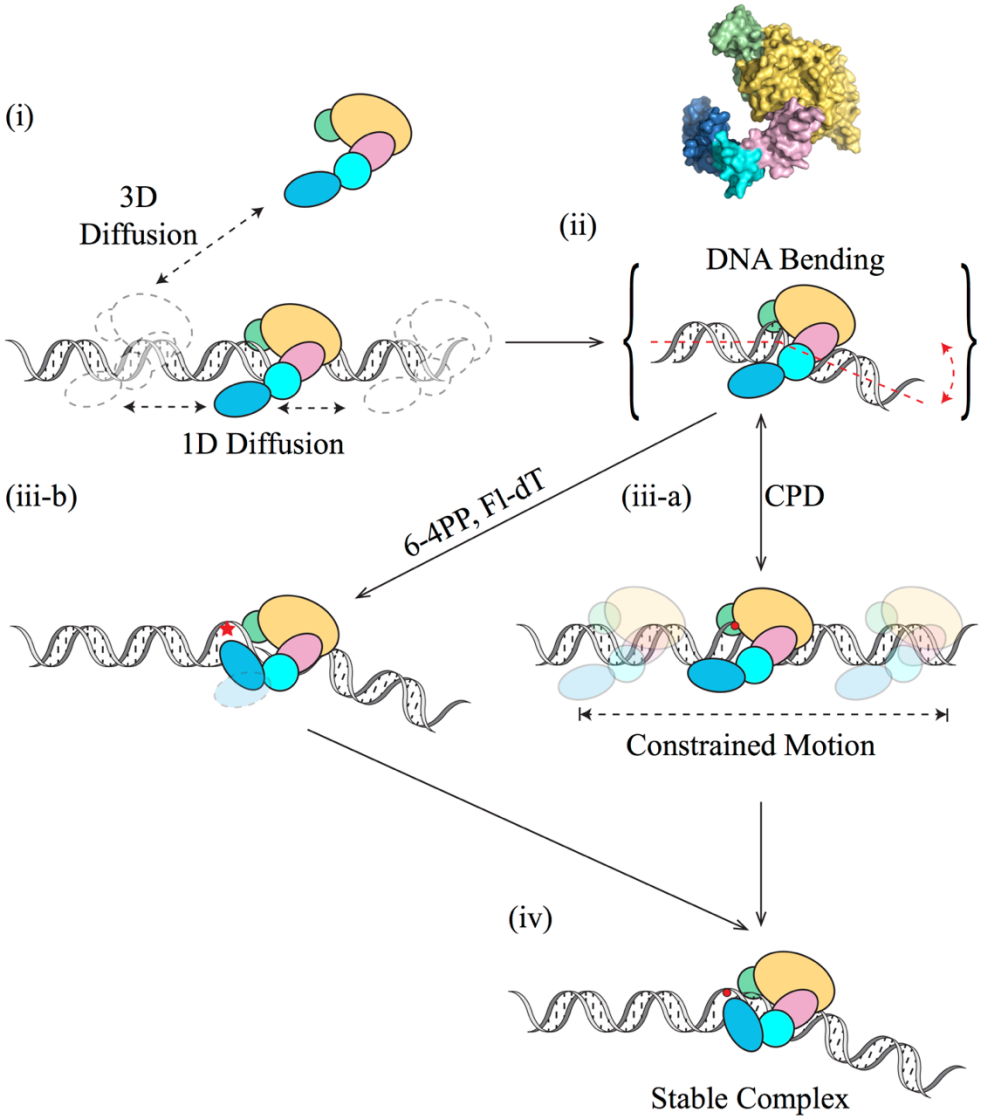

Figure 3.23: Working model for dynamic lesion recognition by Rad4-Rad23.

Domains of Rad4-Rad23 (PDB: 2QSF) color-coded as shown in model: TGD – yellow, BHD1 – pink, BHD2 – cyan, BHD3 – blue, Rad23 – green.
Rad4-Rad23 scans DNA through 3D or 1D diffusion (i) and tests integrity of DNA via bending/twisting during 1D diffusion on DNA (ii). Depending on the type of damage encountered, Rad4-Rad23 can either undergo constrained motion around lesion due to lack of β-hairpin 3 insertion (iiia), or alternatively rapidly forms stable protein-DNA complex at site of lesion with β-hairpin 3 inserted for stabilization in a twist-open action (iiib). While it is possible that the DNA in (iiia) is bent, for simplicity this is not shown. Extra time spent probing the lesion, afforded by constrained motion of Rad4-Rad23, could also lead to stable binding of the protein at sites that require larger base opening/flipping energies (iv). Adapted with permission from (279).

Our working model of recognition suggests that, Rad4-Rad23 scans DNA for lesion through a combination of 3D and 1D diffusion (Figure 3.23(i)). Using β-hairpin domains 1 and 2, Rad4 checks the integrity of DNA through bending or twisting (200), either during random 1D sliding on DNA or when transiently bound to DNA during 3D search (Figure 3.23(ii)). When diffusing one dimensionally on DNA, a protein conformational change could be triggered by lesion encounter, which may allow Rad4 to enter a binding state of stronger interactions with DNA, resulting in the subdiffusion of Rad4-Rad23 on DNA. Both DNA bending/twisting and protein conformational changes could contribute to shorter regions of interrogation by the protein (Figure 3.23(iii-a)). Alternatively, spontaneous helix opening and base flipping, which are energetically linked to DNA bending, may follow at sites of severely helix-destabilizing lesions, facilitating immediate β-hairpin 3 insertion and leading to stable binding (Figure 3.23(iii-b)). Otherwise, BHD3 continues to interact with and probe DNA as the protein undergoes constrained motion. As the subdiffusive protein has limited range, β-hairpin 3 insertion is afforded more opportunities to proceed, therefore leading to recognition of difficult targets with slower base-flipping rates, consistent with the previously published kinetic gating mechanism of Rad4-Rad23 damage recognition (199,200). Both recognition pathways, through rapid spontaneous base flipping or constrained motion, converge to form a stable recognition complex (Figure 3.23(iv)). We hypothesized that the complete damage recognition steps may need to progress through all three stages of protein-DNA interaction, namely from random 1D diffusion to constrained motion, followed by stable binding. However, careful analysis of all kymographs (n ~ 1600) generated for this study did not yield any definitive transitions in either direction between the diffusive modes of the protein. We speculate that both 1D diffusive states are stabilized by binding energy

contributed from protein-DNA interactions and the barrier to transitions are large such that these transitions are rare and rapid, therefore were not captured within the observation time scale.

3.8 CONCLUSION

In summary, using single-molecule fluorescence microscopy, we have shown that Rad4-Rad23 performs both random walk and subdiffusion to facilitate damage recognition at different lesions. Evidence suggests that Rad4 β-hairpin domains 1 & 2 induced DNA bending, independent of β-hairpin domain 3 and thus most likely an early step and allows damage recognition for Fl-dT. Taken together our data support a dynamic multi-step damage recognition model utilizing different structural domains for distinct stages of damage detection.

3.9 ACKNOWLEDGEMENT

The author wish to thank Emily Beckwitt for critically reading and commenting on the manuscript; Dr. Brenda Diergaarde for helping with statistical analyses; Dr. Elise Fouquerel for helping with slot blot assays; and Drs. Guillermo Romero, Marcel Bruchez, and Patricia Opresko for providing helpful discussions. This work was made possible through funding from the NIH 5R01ES019566 to B.V.H., 5R01ES024872 to K.A.B., 5R01ES002614 to J.J.W., 2P30CA047904 to University of Pittsburgh Cancer Institute, and from the National Science Foundation MCB-1412692 to J.-H.M.

4.0 DISCUSSION[3]

In the previous chapter, we characterized the diffusive behavior of quantum dot-labeled Rad4-Rad23, in which the complex used facilitated diffusion to search for lesions in DNA. In addition to random diffusion, we also observed apparent subdiffusive motion of Rad4-Rad23 on DNA, which we termed 'constrained motion'. In the sections below, the physical origins and models of anomalous subdiffusion, pertaining to 1D diffusion of protein on DNA, are deliberated. Future directions including further studies on Rad4-Rad23 and its human homologs XPC-RAD23B, as well as experiments investigating lesion handoff to and from these damage recognition complexes during NER, are also discussed at the end.

4.1 ALTERNATIVE EXPLANATIONS OF APPARENT SUBDIFFUSION

Before concluding that this constrained motion was linked to a pattern of anomalous diffusion, we first had to rule out three alternative scenarios that could also explain the observed behavior. High frequency drift due to system noise was considered first and easily ruled out by observing that particles showing subdiffusive behavior were observed alongside other randomly diffusing or non-

[3] Sections 4.1 – 4.3 are adapted with permission from the following published manuscript: Kong, M. and Van Houten B. (2017) Rad4 recognition-at-a-distance: Physical Basis of conformation-specific anomalous diffusion of DNA repair proteins. *Progress in Biophysics and Molecular Biology*, **127**, 93-104.

motile particles bound to the same DNA tightrope. Additionally, the extent of subdiffusion changed in response to either different substrates or protein constructs, indicating that those factors acting as the sole variable between sets of experiments were the cause of change in constrained motion. Systematic noise coming from the microscope stage or intrinsic to the assay platform would have caused global changes in diffusive behavior regardless of other variables. Secondly, it had been noted that apparent anomalous subdiffusion could arise from errors in determining particle positions in single particle tracking experiments (296). However, given the error in the tightrope platform combined with measured diffusion coefficients, our measurement times (~1000 s) and lengths of traces used in the fitting process (~10 –100 s) were at least an order of magnitude larger than the characteristic time (~0.1 – 1 s), at which point, according to the arguments raised by Martin *et al.*, the measured anomalous diffusion exponent α_{ap} approaches within 90% of true α. Finally, we considered the possibility of highly subdiffusive behavior due to the energetic constraint imposed by protein-induced super-helical torsional stress in DNA. We envisioned a DNA unwinding model similar to that of RNA polymerase, which, as it translocates during transcription, generates positive supercoiled waves in front of itself and negative super-helical stress behind (297). By analogy, we envisioned that β-hairpin 3 of Rad4, which is melted into the DNA in the co-crystal structure (Figure 3.1) (88,199), could remain engaged with DNA during diffusion and cause positive supercoils to build up ahead of the protein in the direction of motion and thus impede further movement. However, our calculations revealed that such a mechanism would only allow Rad4-Rad23 to travel about 50 – 100 bp in either direction with the thermal energy ($k_B T$) at room temperature, an order of magnitude smaller than observed (500 – 1000 bp) and below our criterion for a particle being considered motile (motion of three pixels, or ~500 bp at 46 nm/pixel).

What then is causing this constrained motion of Rad4-Rad23 around photoproducts in DNA? The answer requires a thorough understanding of factors that contribute to anomalous diffusion. The following sections focus on the basis of 1D anomalous diffusion of protein on DNA. For a review of theoretical aspects and physical contributions to anomalous diffusion and examples of such behavior in a wide range of different types of biological systems, please refer to APPENDIX C.

4.2 ONE-DIMENSIONAL (SUB)DIFFUSION OF PROTEIN ON DNA

Given that the two-state model is a generally accepted solution to the speed-stability paradox in target search (see 1.4.2.2), it is important to revisit models of one-dimensional diffusive behavior of a protein on DNA without obstacles. A simple and intuitive way to capture base-sequence-dependent protein-DNA interactions was derived and used to model nonspecific one-dimensional sliding on DNA by Barbi *et al.* (298,299). A model was constructed based on the idea that a sequence-specific protein "reads" the underlying sequence from the DNA major groove while sliding and that recognition is achieved by formation of a specific set of hydrogen bonds between the protein amino acids and the target sequence bases. This approach also assumes that the protein attempts to make the same set of hydrogen bonds on nonspecific sequences as it does at target sites. Protein-DNA interaction at position *n* was expressed as a $4 \times m$ matrix (D_n), with *m* being the size of the recognition sequence. The recognition matrix (R) can also be constructed, based on known protein-DNA contacts from structural data, as an $m \times 4$ matrix. The interaction energy landscape at base *n*, with the implicit assumption that energy contributions from hydrogen bonds are additive, is thus defined as

$$E(n) \propto tr(R \cdot D_n) \tag{4.1}$$

A case study of T7 RNA polymerase promoter search on T7 DNA with different translocation mechanisms, including one variant of the two-state model, was conducted through simulations. Diffusion was found to be anomalous and subdiffusive for short times and asymptotically approached normal over longer time periods (299). Furthermore, it was shown that this formulation with energy contributions from discrete hydrogen bonding events could be generalized, which led to Gaussian-distributed interaction energies ($\sigma \sim 2.5\ k_B T$) and gave rise to quantitatively similar one-dimensional diffusive behavior as before (298). Similar transient anomalous subdiffusion due to trapping effects was obtained through another set of Monte Carlo simulations, where diffusion on DNA was modeled as a random walk on a one-dimensional lattice with different models for traps (300). In such a system, an infinite hierarchy of traps is believed to lead to subdiffusion through a CTRW mechanism. Nonetheless, it was shown that in the generalized case of finite binding site hierarchy, where the target site was represented by the deepest trap, random energy model with a continuous Gaussian distribution ($\sigma = 1.5\ k_B T$) recapitulated the transient nature of subdiffusion and its crossover to normal diffusive behavior (300).

At its roots, the speed-stability paradox and its solution are connected to the fact that the one-dimensional diffusive behavior of a protein on DNA is affected by a random potential landscape (226). For a one-dimensional diffusing particle in a random potential with Gaussian-distributed amplitudes, its diffusion coefficient is proportional to $e^{-\sigma^2}$, where σ^2 denotes the variance of the Gaussian distribution (301), as discussed earlier in the section. The use of a random Gaussian-distributed potential as a continuum approximation of the nonspecific sequence-dependent interaction between a DNA binding protein (e.g. a transcription factor) and the DNA

sequence being scanned can be justified by the overall heterogeneity in nucleotide sequence for sufficiently long regions (302). Inspired by the experimental evidence that binding of the *Cro* repressor induced bending at both specific and nonspecific sites (303), Mirny and colleagues sought to refine the description of the random energy landscape used in the modeling of one-dimensional diffusion. They argued that there exists a finite-range correlation, whose length scale is on the order of the size of the protein binding domain (304). It was shown that diffusion in a correlated potential is slower than in an uncorrelated potential, and that the mean first passage time (MFPT) fluctuates more in a correlated potential. When the length scale of diffusion is less than the characteristic distance (N_c), where, by definition, there is no self-averaging, subdiffusive as well as superdiffusive behavior can occur. Simulations of random walks suggested that as a result of the correlated random potential, proteins could preferentially localize in certain areas of the genome. Diffusion on a correlated random potential was recently revisited by Goychuk and Kharchenko (305). They reasoned that the interaction energies between each base in contact with the protein are additive and that spatial correlation arises because when the protein slides by one base on DNA, the same set of bases remains in contact with the protein, except for the one farthest away from the direction of the movement. Starting with the Langevin equation and the assumption for an exponentially decaying short-range correlation, Goychuk and Kharchenko first showed that such correlation has no effect on the scaling of the diffusion coefficient and that the corresponding diffusion is ergodic in the macroscopic scale. An equation for the mesoscopic subdiffusion was then derived to estimate the physical length scale at which subdiffusion would be expected due to correlations in potential energy. Subdiffusion was shown to be readily macroscopic for a Gaussian potential energy disorder $\sigma \sim 4 - 5\ k_B T$. Surprisingly, target site location via such subdiffusion was also shown to proceed faster than expected.

4.3 CONFORMATION-DRIVEN CONSTRAINED MOTION OF RAD4-RAD23

As discussed above, sequence-specific DNA binding proteins, such as transcription factors and restriction enzymes, have been at the center of many studies on the theoretical modeling of target search. However, parallels can be drawn to proteins that recognize other features of DNA. Such is the case for a wide range of DNA damage sensing proteins. In addition to DNA binding proteins like TRF1 (292) and the stromal antigen subunit (SA) SA1 (306), DNA repair proteins such as DNA glycosylases (291), UV-DDB (43), and endonuclease UvrC from bacterial NER (266) have all been observed to exhibit subdiffusion to some extent. Although not explicitly tested, transient subdiffusion may also be a property of the eukaryotic mismatch repair complex Msh2-Msh6 (261) and restriction enzyme EcoRI (307).

Based on our findings on target search by Rad4-Rad23, we propose that one-dimensional constrained motion of proteins on DNA may result from diffusion in a potential energy landscape due to extended protein-DNA interactions and may be a functional form of target recognition *in vivo*. Notably, factors that affect the extent of observed constrained motion include ionic strength of the solution, type of lesion in DNA, and the presence of β-hairpin 3 of Rad4. Increased constrained motion as a result of the loss of the residues located at the tip of β-hairpin 3 appeared to exhibit a compensatory effect as the protein remained biologically functional *in vivo*; neither UV resistance nor photoproduct repair was compromised in yeast expressing this mutant. Metzler and colleagues suggested that three-dimensional subdiffusion of transcription factors helps to keep them in the vicinity of their targeted binding sites in DNA and may be beneficial to gene regulation *in vivo* (308). We proposed constrained motion by a protein around its target site as a mechanism for “recognition-at-a-distance.” Rad4’s ability to participate in productive NER while undergoing such constrained motion can be thought of as a first responder to arrive at the scene of an accident,

namely the ability to direct other emergency personnel around the scene without being physically stationed there at all times. Similar to the *Cro* repressor, we found, using atomic force microscopy, that Rad4-Rad23 bends DNA to ~42° at both specific and non-specific sites. Applying our estimated roughness of diffusional energy landscape and footprint of Rad4 on DNA based on the co-crystal structure to the one-dimensional subdiffusion as modeled in 4.2 (304,305), subdiffusion may be expected to emerge on the length scale of ~400 – 800 bp, roughly consistent with the observed range of constrained motion (~500 – 1000 bp). Furthermore, we consider the extent of such subdiffusive behavior to be linked to the specific conformation that the protein adopts while interacting with DNA. The strength of correlation in the protein-DNA interaction potential may be influenced by the structural motif(s) that are probing the underlying sequence and the structural integrity of the sequence itself. Most importantly, this one-dimensional subdiffusion, driven by the specific conformation adopted by the protein-DNA complex in general, may represent one intermediate in a generalized two-state model (Figure 4.1). In the case of Rad4, the protein interacts with UV-irradiated DNA and was observed to form: 1) molecules that show random linear motion with a low barrier of diffusion of $\sim 0.4\ k_B T$ and an anomalous diffusion exponent $\alpha \sim 1$ (Figure 4.1(i)); 2) molecules showing constrained motion with a barrier to free diffusion on DNA of about $\sim 1.6\ k_B T$ and an anomalous diffusion exponent $\alpha < 0.8$ (Figure 4.1(ii)); and finally, 3) non-motile complexes that we believe represent stable specific binding complexes (Figure 4.1(iii)). Since a mutant of Rad4 lacking the entire β-hairpin domain 3 was able to bind specifically to sites of damage and bend the DNA at sites of damage by ~37°, β-hairpin domains 1 and 2 of Rad 4 must make large contact with DNA, and are capable by themselves of transiently bending the DNA (Figure 4.1(ii)). Presumably this is mediated by the β-hairpin of domain 2. This protein-induced bend would help to increase the energy landscape of the DNA (green) and favor DNA opening,

producing a sufficiently steep landscape that can induce constrained motion with an anomalous diffusion exponent $\alpha < 1$. In this manner, engagement of additional structural motifs on the target recognition path, which goes from freely diffusing on DNA to stably bound recognition complex, could constitute as different intermediates that correspond to increasing levels of ruggedness in the diffusional energy landscape. Thus the correlated potential energy profile from protein-induced DNA bending gives rise to the observed constrained motion. In fact, molecular dynamics simulations and measurements of the free-energy path of Rad4-Rad23 interaction with a mismatched CPD indicated that Rad4's interaction proceeds via an induced fit model, rather than a structural capture model (201). This idea is also consistent with the notion that UV-DDB interacts with damaged DNA using a conformational proofreading mechanism (43).

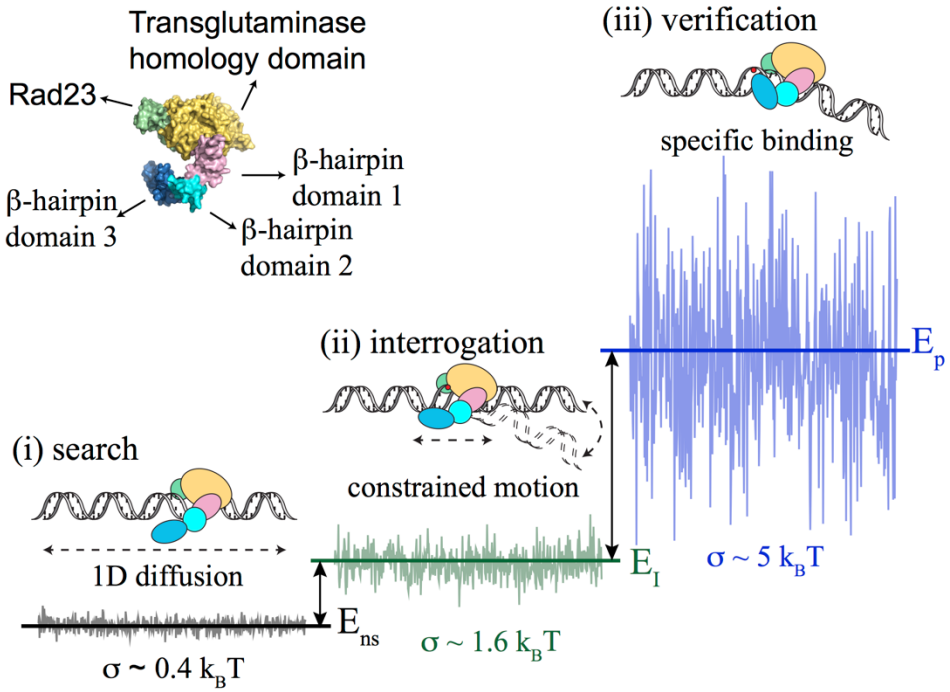

Figure 4.1: Model for conformation-driven constrained motion of Rad4-Rad23.

Domains of Rad4 are as colored and labeled.
(i) Rad4-Rad23 diffuses randomly on DNA where non-specific protein-DNA interactions contribute to the smooth energy landscape ($\sigma \sim 0.4\ k_BT$).
(ii) Interrogation of DNA through interactions with β-hairpin domains 1 and 2, that most likely includes transient DNA bending, coupled with correlations in potential energy along the DNA due to the presence of lesions, lead to increased ruggedness in the energy landscape ($\sigma \sim 1.6\ k_BT$). This constrained motion and subdiffusive behavior emerges on the scale of 500-1000 bp, and may represent 'recognition-at-a-distance.'
(iii) Specific damage verification and binding is achieved through β-hairpin 3 insertion, which results in a much rougher energy landscape($\sigma > 5\ k_BT$). The insertion step occurs spontaneously at highly helix-distorting lesions, or is facilitated when Rad4 is slowed down while undergoing constrained motion near the damage site.
Adapted with permission from (206).

Again, the readers are encouraged to visualize the recognition-at-a-distance model as akin to what happens at the scene following a traffic accident. A law enforcement officer (Rad4-Rad23), who typically patrols the streets (search for DNA damage via 3D or 1D diffusion) looking for any type of traffic violations (DNA quality check via bending), arrives at the scene of an accident (site of DNA lesion). In a horrific crash (a severely helix distorting lesion, such as 6-4PP), the officer

immediately checks on the injured (stable binding at site of lesion). While attending to the seriously wounded, the officer radios for further assistance from emergency medical services (recruitment of downstream repair factors). During the process, the officer is extensively involved in providing immediate aid at the scene such that (s)he remains engaged (stable binding of Rad4-Rad23 due to rugged energy landscape as a result of β-hairpin 3 insertion). In comparison, should the officer encounter a relatively minor fender bender (less helix distorting lesion, such as CPD), after initially checking on those involved, (s)he is able to remain around the scene (subdiffusion around lesion) to direct traffic and contacting dispatch for help (signaling for DNA damage). Even though the officer is not bound to the scene, (s)he is able to communicate with the people involved while coordinating the arrival of other emergency responders (continued probing of lesion, with a smoother energy landscape, during subdiffusion). Finally, perhaps after setting up traffic cones in the vicinity, the officer returns to the scene to conduct interviews and write up reports (stable binding at site of lesion after a period of subdiffusion). For an observer of these events (experimental observations), we may not witness every stage of the response as we pass by the scene (lack of transition between modes of diffusion during the observation window). Although it could be argued that given longer observation periods or the ability to follow the officer around (longer experiments or continuous tracking of particles), we should be able to watch the officer go through each step of the process.

Despite the wealth of experimental data on anomalous subdiffusion in diverse biological systems, the underlying physical mechanism of such behavior has yet to be fully elucidated. In particular, subdiffusive motion exhibited by proteins while sliding on DNA during target search has not been examined as closely compared to various models of three-dimensional subdiffusion, theoretically or experimentally. Better modeling of the physical basis of such behavior could

contribute to greater characterization and understanding of biological systems involving sequence- or structure-specific DNA binding proteins, as well as more precise tuning of known protein-DNA interactions in engineered systems. Advances in imaging techniques and computing technology, single-molecule/single-cell experiments, and simulations based on atomic details of proteins and base-sequence of DNA could provide key insights into solving the puzzle.

4.4 FUTURE WORK

Based on our proposed 'recognition-at-a-distance' model of damage sensing and its biological relevance as observed in UV survival studies, as well as the general behavior of Rad-Rad23 characterized in this dissertation, a series of experiments with testable hypothesis can be designed and conducted to further improve our understanding of damage recognition and lesion handoff in the repair process. Conclusions from this and future studies can also be potentially applied to other processes that involve multiple protein factor assemblies at target sites on DNA.

1. *Further dissection of the structure-function relation:* Crystal structure indicates that TGD and BHD1 contact only the undamaged DNA duplex. This leads to the simple hypothesis that the protein construct consisting of only TGD and BHD1, should it be stable, would not exhibit damage specific binding and would diffuse randomly on DNA. MD studies have shown that initial engagement of BHD2 with the minor groove of DNA leads to further DNA bending (201). DNA bending at specific and nonspecific sites in DNA maybe dependent on the presence of BHD2. Since damage recognition is not totally abolished in BHD3 deletion mutant, it is reasonable to hypothesize that BHD2 carries out part of the damage sensing function through DNA bending. Using a ΔBHD2 mutant, it may be

possible to decouple the DNA bending component from damage-specific. Furthermore, knowing that the 'base-flipping highway' involves residues F556, F597, and F599, biochemical and single-molecule characterizations of the F556A/F597A/F599A triple mutant would underscore the importance of these residues.

2. *Examination of the lesion structure dependence:* As we have established in this study using Fl-dT and CPD lesions, damage recognition may progress through different paths depending on the type of lesion being examined. This idea is supported by a recent MD simulations study on recognition of benzo[α]pyrene adducts by Rad4-Rad23 (203). Distinct chemical structures of lesions also differentially influence the local conformation and flexibility of DNA, which in turn facilitate DNA bending by β-hairpin 2 and the insertion of β-hairpin 3. Using defined lesion substrates, various NER substrates can be incorporated to test the hypothesis that chemical structures of lesions could dictate how they are captured and recognized by Rad4-Rad23. Future studies on the subject of lesion dependence in damage recognition by XPC-RAD23B may also include oxidative DNA lesions. XPC has been shown to not only be involved in BER (309,310) but also bind to 8-oxo-7,8dihydroguanine (8-oxoG) (311), allowing simulation of the OGG1 activity (309).
3. *Delineation of the parallels between Rad4-Rad23 and XPC-RAD23B:* A natural extension of this work is to characterize damage recognition by the human proteins XPC-RAD23B. Although both Rad4-Rad23 and XPC-RAD23B are required for CPD removal in *S. cerevisiae* and humans, respectively, recognition of CPDs in the latter has been attributed to UV-DDB, which binds CPD with ~10 nM affinity (42), while XPC-RAD23B does not distinguish CPD from undamaged dsDNA (78). In mammalian cells, the requirement of XPC-RAD23B at CPD sites during GG-NER is based on its ability to recruit TFIIH

through its C-terminal residues. The possibility remains, however, that XPC-RAD23B interacts with CPDs in a transient manner that precludes observations in bulk experiments. Furthermore, there exist pathogenic and non-pathogenic point mutations in XPC, whose effects on search mechanism and damage recognition warrant further investigations using single-molecule techniques and comparisons to those in Rad4-Rad23. A tryptophan to serine mutation of the evolutionarily conserved residue W690 (W496 in Rad4), W690S, was identified in XP-C patient XP13PV (177). W690S is known to destabilize XPC and decrease its affinity towards damaged and undamaged DNA *in vivo* and *in vitro*, while retaining its interactions with other NER factors (117,263,312,313). Potentially, XPC W690S could slide on damaged DNA, much the same way that DDB2 K244E was found to behave (43). Another mutation, E755K in XPC, within the so-called 'β-turn' region, was described in a study to reduce nuclear mobility of the protein (264). This glutamic acid, E755, is conserved in higher eukaryotes and maps to T555 in Rad4. Although T555 does not make contact with the DNA itself, the residue is curiously situated in between N554 and F556, both of which interact with and stabilize the 3' flipped-out thymidine in the structure (88). The effects of E755K may manifest as both increased percentage of nonmotile particles on undamaged DNA and lower values of diffusion coefficients of those mutant that do undergo 1D diffusion. For the three most recently identified pathogenic mutations, their corresponding residues in Rad4 do not interact with DNA according to the crystal structure: Y585C is conserved in Rad4 Y379, located in the TGD region on the inside of the protein (175); P703L corresponds to C509 located in the stem of the Rad4 β-hairpin 2 (178); and while T738A is also conserved in Rad4, T537 is in the region connecting BHD2 and BHD3 (179). These mutations have all been proposed to destabilize

XPC but their effects on DNA binding and damage recognition have not been investigated. Lastly, a prior study of point mutations on the conserved residues F756, F797, and F799 in human XPC-RAD23B (F556, F597, and F599 in Rad4) found that these mutations resulted in varying degrees of diminished DNA binding *in vitro* and impaired localization as well as TFIIH recruitment at UV-damaged sites *in vivo* (314). Analyses on these XPC point mutants can also be compared to those equivalent in Rad4.

4. *Characterization of the effects of binding partners and post-translational modifications on damage recognition:* Both CETN2 and Rad33 are known to stimulate NER by forming complexes with XPC-RAD23B and Rad4-Rad23, respectively. CENT2 is also known to stimulate ssDNA binding by XPC (117). Therefore, trimeric complexes XPC-RAD23B-CETN2 and Rad4-Rad23-Rad33 should exhibit better damage recognition both on DNA tightropes as well as in AFM studies. It is possible that the increased binding affinity could only be observed on less distorting lesions, such as CPD, which are otherwise less efficiently recognized. An alternative hypothesis is that the presence of CETN2 and Rad33 in these complexes allow them to better distinguish undamaged DNA from damaged DNA, *i.e.* higher specificity for damaged DNA. Preliminary EMSA-based competition experiments using Fl-dT modified dsDNA as substrate and undamaged dsDNA as competitor have shown that more Rad4-Rad23-Rad33 remain associated with the substrate than Rad4-Rad23 at the same concentration of competitor DNA (unpublished data, Ananya Mukundan and Muwen Kong). Another factor that influences DNA binding affinity of XPC-RAD23B is its status of post-translational modifications. Poly-ubiquitination of XPC by the UV-DDB-containing E3 ubiquitin ligase in the presence of 6-4PP was first shown to increase the affinity of XPC-RAD23B for damage binding *in vitro* (51). Preliminary

evidence also suggests that Ser350 and Ser892 of XPC may be phosphorylated by ATM and ATR in response to DNA damage (315), and that dephosphorylation of Ser892 by wild-type p53-induced phosphatase 1 (WIP1) may lead to inactivation of XPC (316). PARylation of both subunits of XPC-RAD23B, which can be achieved in a UV-dependent manner *in vitro*, has been shown to weaken its interaction with DNA (91). Quantitative changes in distributions of motion types on damaged and undamaged DNA, as well as distributions of observables in AFM experiments, are expected.

5. *Visualization of damage handoff to and from XPC-RAD23B:* A key implication of the proposed 'recognition-at-a-distance' model is that damage signaling and handoff can be accomplished while the damage sensor is undergoing subdiffusion and not stably bound. But does handoff from UV-DDB to XPC-RAD23B still allow the latter to diffuse around the lesion, given that UV-DDB binds stably to damage? How efficient is recruitment of TFIIH by XPC-RAD23B that is stably bound compared to one undergoing constrained motion? How does motion by XPC-RAD23B on DNA factor in the *in vitro* reconstituted NER of CPD, where the process is known to not require UV-DDB but rather proceed through coordination between XPC, TFIIH, XPA, and RPA (317)? P334H, the first identified pathogenic missense mutation in XPC, has been shown to impair interactions with XPA, TFIIH, and BER-related glycosylase OGG1 (99). The damage handoff process to and from XPC-RAD23B can be influenced by post-translational modifications on the protein as well. Firstly, SUMOylation of XPC promotes lesion handoff from UV-DDB to XPC-RAD23B (94). Then poly-ubiquitination of XPC by UV-DDB and RNF111, the latter targeted to XPC via SUMOylation, helps release XPC from damage site after TFIIH recruitment (93). Finally, for *in vitro* studies that use naked DNA substrates, how damage

recognition is carried out in the context of chromatin is one of the most frequently asked questions. The presence of nucleosomal structure in damaged DNA is known to significantly inhibit of reconstituted NER *in vitro* (318). Even though UV-DDB readily binds to 6-4PP-containing nucleosomes regardless of the orientation of the lesion (319), loading of XPC-RAD23B in a reconstituted single-molecule reaction likely still requires chromatin remodeling, such as through the actions of ATP-dependent SWI/SNF (320) or ACF (321), among other mechanisms (322).

4.5 CONCLUDING REMARKS

This dissertation characterizes the NER damage recognition process by Rad4-Rad23 and its structure-function relations through bulk biochemistry and single-molecule biophysics. The observation of lesion- and protein-dependent subdiffusive motion *in vitro*, its biological relevance in cell survival and repair kinetics, as well as the generalized 'recognition-at-a-distance' model and its implications in multi-protein assemblies on DNA are the most intriguing aspects of this work. While the physical origins of subdiffusive motion of proteins on DNA remains curiously elusive, we hope that our discussion and model of conformation-driven constrained motion would encourage future theoretical and experimental studies on this topic.

APPENDIX A

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

AAF – 2-Acetylaminofluorene

AFM – Atomic force microscopy

BER – Base excision repair

bp – Base pair

BPDE – Benzo[α]pyrene diol epoxide

CAK – CDK-activating kinase

COP9 – Constitutive photomorphogenesis 9

CPD – Cyclobutane pyrimidine dimer

$CRL4^{CSA}$ – DDB1-CSA-CUL4-RBX1 cullin-RING E3 ubiquitin ligase

$CRL4^{DDB2}$ – DDB1-DDB2-CUL4-RBX1 cullin-RING E3 ubiquitin ligase

CS – Cockayne syndrome

CSA – Cockayne syndrome WD repeat protein A

CSB – Cockayne syndrome B protein

CSN – COP9 signalosome

CTRW – Continuous-time random walk

DDB1 – DNA damage-binding protein 1

DDB2 – DNA damage-binding protein 2

E. coli – Escherichia coli

FBM – Fractional Brownian motion

FGN – Fractional Gaussian noise

FLE – Fractional Langevin equation

GG-NER – Global genome nucleotide excision repair

HDP – Heterogeneous diffusion process

HMGN1 – High mobility group nucleosome-binding domain-containing protein 1

MMR – Mismatch repair

MSD – Mean squared displacement

NER – Nucleotide excision repair

NMSC – Non-melanoma skin cancer

nt – Nucleotide

PARP1 – Poly(ADP-ribose) polymerase 1

PARylation – Poly-ADP-ribosylation

PCNA – Proliferating cell nuclear antigen

QD – quantum dot

Rbx1 – E3 ubiquitin-protein ligase

RFC – Replication factor C

Roc1 – Regulator of cullins 1

S. cerevisiae – Saccharomyces cerevisiae

SAQD – Streptavidin-conjugated quantum dot

SBM – Scaled Brownian motion

ssDNA – Single-stranded DNA

TC-NER – Transcription-coupled nucleotide excision repair

TFIIH – Transcription factor IIH

TFIIS – Transcription factor IIS

TTD - Trichothiodystrophy

TTDA – Trichothiodystrophy protein A

TTDN1 – TTD non-photosensitive 1 protein

USP7 – Ubiquitin carboxyl-terminal hydrolase 7

UV-DDB – UV-damaged DNA-binding protein

UV – Ultraviolet

$UV^{S}S$ – UV-sensitive syndrome

UVSSA – UV-stimulated scaffold protein A

WIP1 – Wild-type p53-induced phosphatase 1

XAB2 – XPA-binding protein 2

XP – Xeroderma pigmentosum

XP-A – Xeroderma pigmentosum, complementation group A

XPA – DNA repair protein complementing XP-A cells

XP-B – Xeroderma pigmentosum, complementation group B

XPB – TFIIH basal transcription factor complex helicase XPB subunit

XP-C – Xeroderma pigmentosum, complementation group C

XPC – DNA repair protein complementing XP-C cells

XP-D – Xeroderma pigmentosum, complementation group D

XPD – TFIIH basal transcription factor complex helicase XPD subunit

XP-E – Xeroderma pigmentosum, complementation group E

XP-F – Xeroderma pigmentosum, complementation group F

XPF – DNA repair endonuclease F

XP-G – Xeroderma pigmentosum, complementation group G

XPG – DNA repair protein complementing XP-G cells

XP-V – Xeroderma pigmentosum variant

APPENDIX B

CHAPTER 3 SUPPLEMENTAL DATA[4]

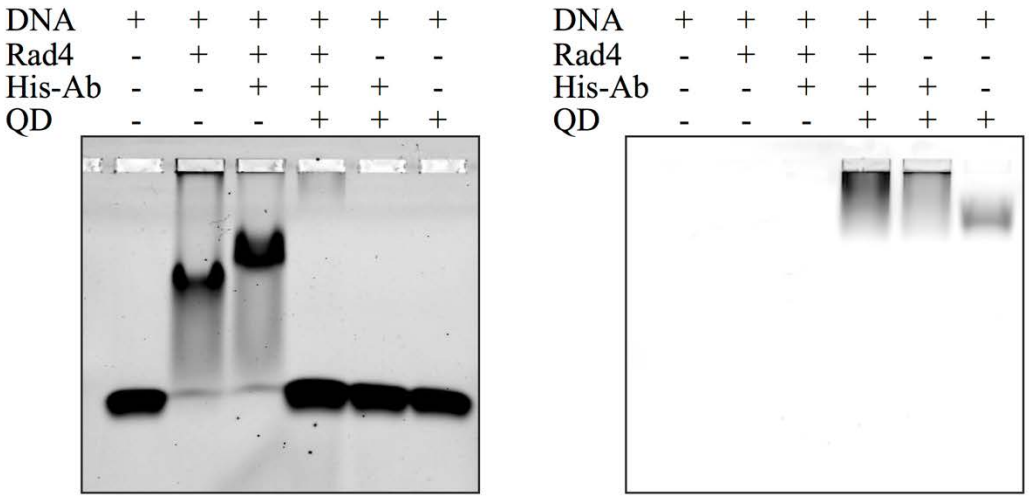

Figure B.1: Electrophoretic mobility shift assay of Rad4-Rad23 conjugated with anti-His antibody and quantum dots.

Agarose EMSA gel of Rad4-Rad23 conjugated with anti-His antibody and quantum dots binding to 37 bp Fl-dT-containing DNA. Imaged using 526 nm (*left*) and 670 nm (*right*) filters, respectively.

[4] This appendix is adapted with permission from the Supplemental Information from the following published manuscript: Kong, M., *et al.* (2016) Single-Molecule Imaging Reveals that Rad4 Employs a Dynamic DNA Damage Recognition Process. *Molecular Cell*, **64**, 376-387.

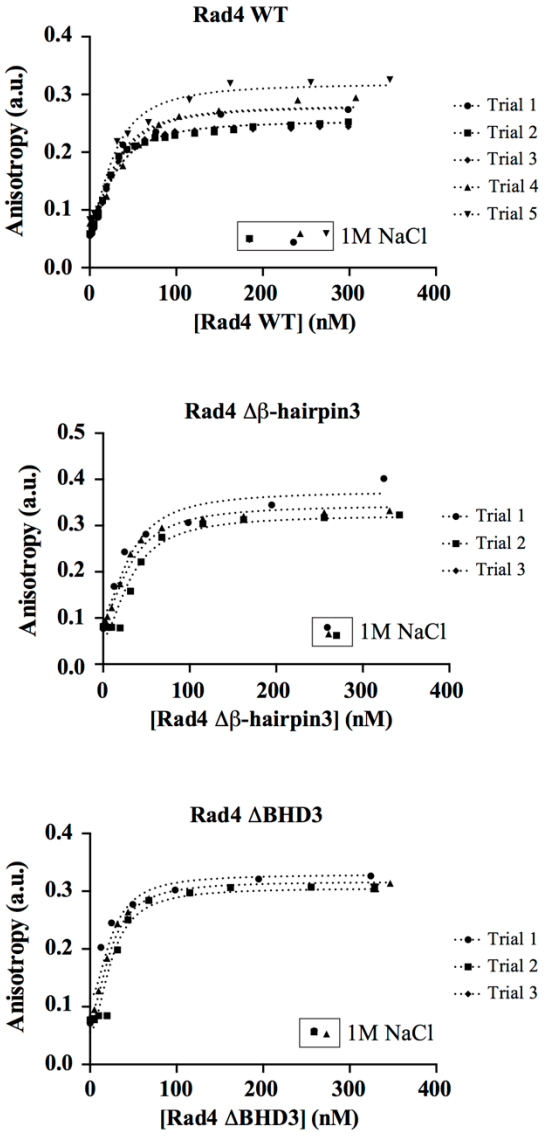

Figure B.2: Steady state binding of Rad4-Rad23 to fluorescein-containing duplex DNA, measured by fluorescence anisotropy.

Equilibrium binding experiments of WT (*top*), Δβ-hairpin3 (*middle*), and ΔBHD3 (*bottom*) were repeated 3-5 times. Dotted curves shown were obtained by fitting replicates from each protein construct globally using a single-site cooperative binding model. Data points in boxes show reversibility of each binding reaction after addition of 1M NaCl.

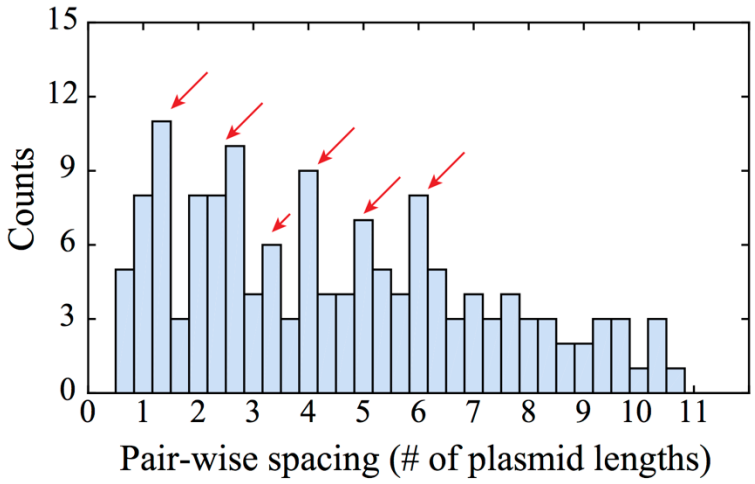

Figure B.3: Distribution of pair-wise distances between stably bound Rad-Rad23 particles on Fl-dT DNA damage arrays.

Measured in numbers of plasmid lengths (2030 bp). Red arrows indicate regular integer multiples of up to 6 plasmid lengths.

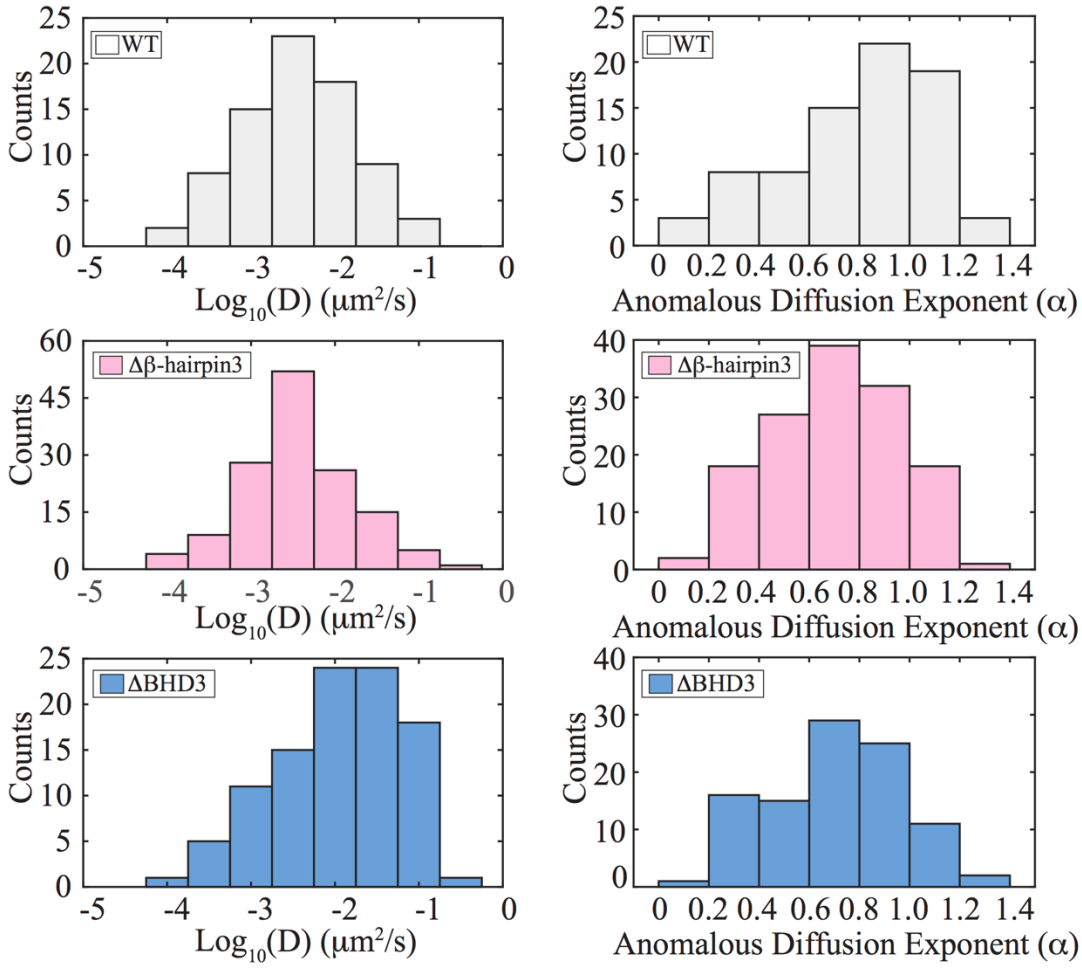

Figure B.4: Diffusive behavior of Rad4 WT and deletion mutants on 20J/m^2 UV-irradiated λ-DNA.

Distributions of diffusion coefficients $log_{10}D$ (*left column*) and anomalous diffusion exponent α (*right column*) of WT Rad4-Rad23 (*top*), Δβ-hairpin3 (*middle*), and ΔBHD3 (*bottom*).

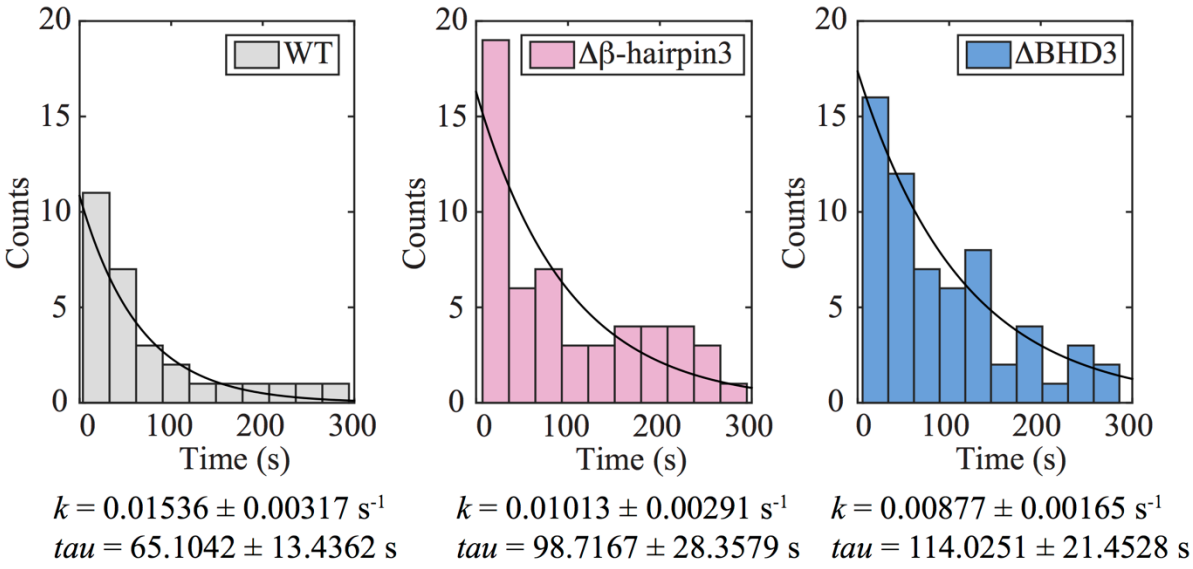

Figure B.5: Rad4 WT and deletion mutants share similar dissociation kinetics.

Single exponential fittings of dwell time histograms of dissociating WT Rad4-Rad23 (*left*), Δβ-hairpin3 (*middle*), and ΔBHD3 (*right*) particles.

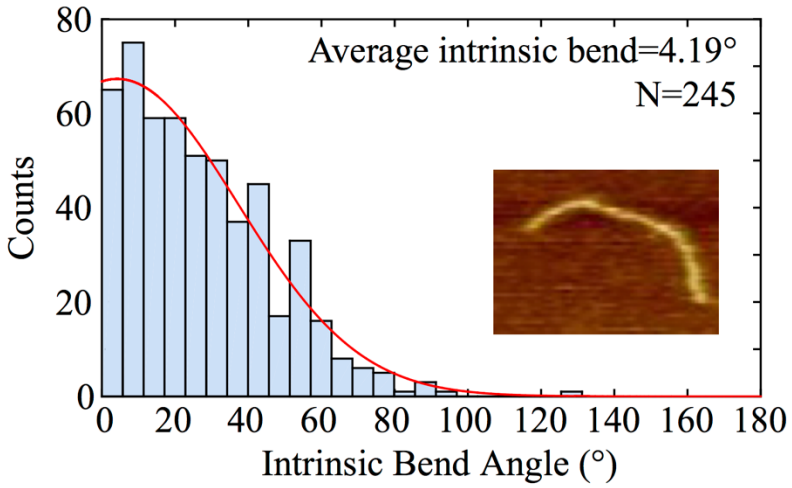

Figure B.6: Intrinsic bend in Fl-dT-containing duplex DNA.

Distribution and Gaussian fitting of intrinsic bend angles of the 538 bp Fl-dT-containing DNA fragment used in AFM experiments. *Inset:* representative AFM image of a 538 bp DNA fragment.

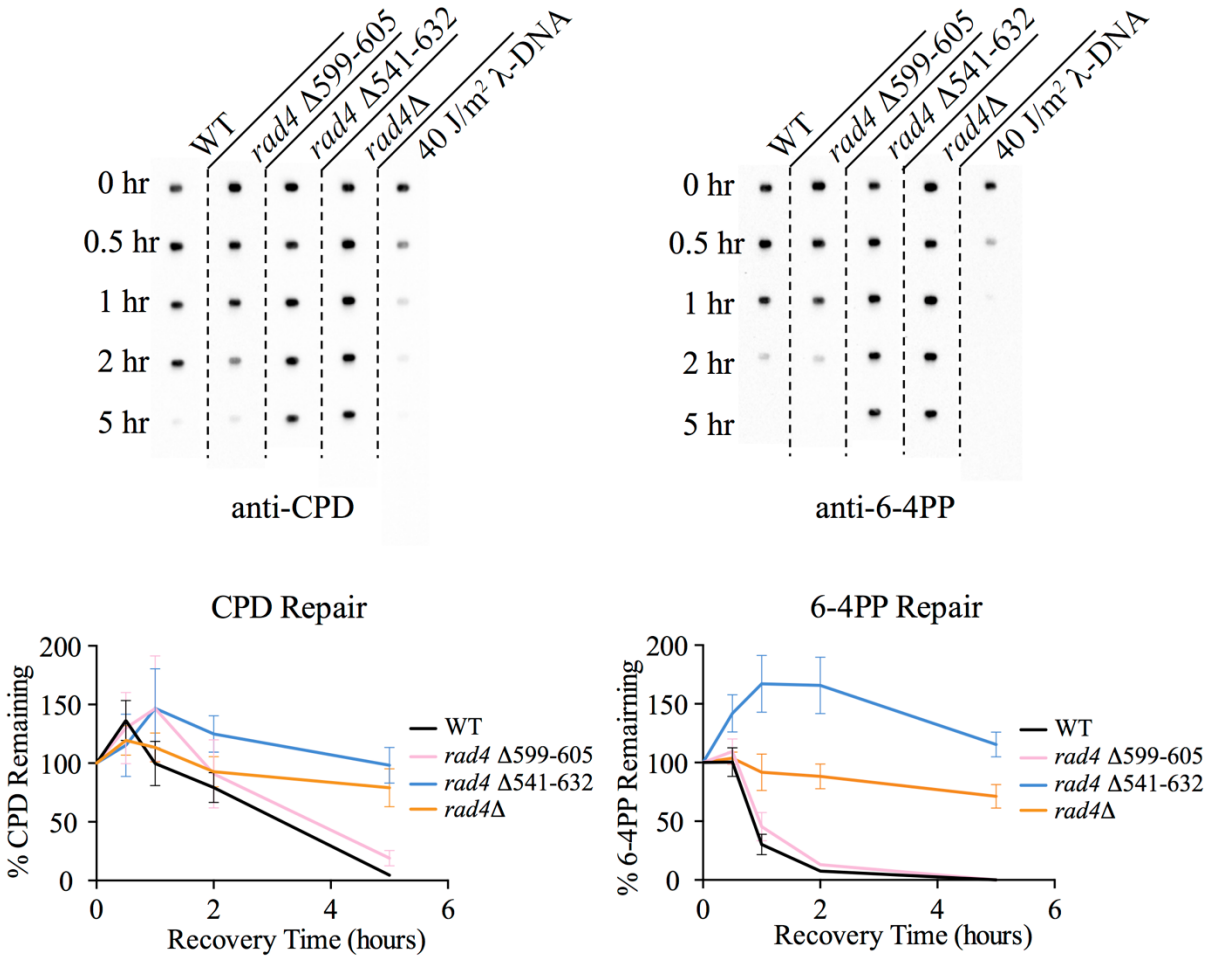

Figure B.7: Antibody slot blots of CPD and 6-4PP repair kinetics.

Top: Representative antibody slot blots against CPD (*left*) and 6-4PP (*right*) in *S. cerevisiae* genomic DNA, post-100J/m^2 UV treatment, with indicated recovery times.
Bottom: Quantitative repair kinetics of CPD (*left*) and 6-4PP (*right*). Time courses shown are averages of two or three biological repeats, each spotted in triplicates. Error bars represent standard errors of the mean of two to three experiments, each done in triplicates.

	WT	Δβ-hairpin3	ΔBHD3
K_D (nM)	28.72 ± 0.65	33.07 ± 1.81	24.89 ± 1.35
Hill coefficient	1.66 ± 0.06	1.72 ± 0.16	1.90 ± 0.18

Table B.1: WT and deletion mutants of Rad4-Rad23 bind tightly to Fl-dT-containing duplex DNA.

Equilibrium dissociation constants (K_D) and Hill coefficients obtained from fitting anisotropy data (3-5 duplicates) with the single-site specific binding with Hill slope model in *Prism 6* (best-fit value ± standard error).

Name	Description	Sequence
OMP007	sgRNA targeting *RAD4*	GATCACAGTTAAGCCAGTTTTAAGGTTTTAGAGCTAG[5]
OMP008	sgRNA targeting *RAD4* (other strand)	CTAGCTCTAAAACCTTAAAACTGGCTTAACTGT
OMP011	Donor DNA for *rad4-Δ599-605* (Forward)	AGGTTTTTGGGGGTGGAATTTGCACCTGCTGTAACTTCTTTTA AGAAGCCAGTTTTAAGTGGCATTGTTGTTGCAAAGTGGCTCA GAGAA
OMP012	Donor DNA for *rad4-Δ599-605* (Reverse)	TTCTCTGAGCCACTTTGCAACAACAATGCCACTTAAAACTGGC TTCTTAAAAGAAGTTACAGCAGGTGCAAATTCCACCCCCAAA AACCT
OMP013	Donor DNA for *rad4-Δ590-615* (Forward)	GAGAATCCTGTGGCAATTAAAGCTGCTAGGTTTTTGGGGGTG GAAAAGTGGCTCAGAGAAGCTATTGAAACCGCTATTGATGGA ATAGAG
OMP014	Donor DNA for *rad4-Δ590-615* (Reverse)	CTCTATTCCATCAATAGCGGTTTCAATAGCTTCTCTGAGCCAC TTTTCCACCCCCAAAAACCTAGCAGCTTTAATTGCCACAGGAT TCTC
OMP015	Donor DNA for *rad4-Δ541-632* (Forward)	GAAGAAGATGAAAGATTATATAGCTTTGAAGACACAGAATTA TACCAAGAGGATGATAATAGGAAGGAACATTTGCTTGGTGCT TTGGAG
OMP016	Donor DNA for *rad4-Δ541-632* (Reverse)	CTCCAAAGCACCAAGCAAATGTTCCTTCCTATTATCATCCTCT TGGTATAATTCTGTGTCTTCAAAGCTATATAATCTTTCATCTTC TTC
OMP017	Donor DNA for *rad4-Δ541-Cterm*[6] (Forward)	GAAGAAGATGAAAGATTATATAGCTTTGAAGACACAGAATTA TACAATGAGGCTGAAACGGTTTGAATAATTAGGAAAGTATGT TTTTAA
OMP018	Donor DNA for *rad4-Δ541-Cterm*[6] (Reverse)	TTAAAAACATACTTTCCTAATTATTCAAACCGTTTCAGCCTCA TTGTATAATTCTGTGTCTTCAAAGCTATATAATCTTTCATCTTC TTC
OMP019	*rad4* screening primer (Forward)	AGGACAGTTGGAAGGCCTAA
OMP020	*rad4* screening primer (Reverse)	GTAGCACTTTCCTCCGCTT
OMP021	*rad4* screening primer (Reverse, for C-term deletion)	ACTCAAGTCCCTGTCCCTCT

Table B.2: Oligonucleotides used for constructing rad4 mutants using the CRISPR/Cas9 system.

[5] The 20mer guide sequence is underlined

[6] The *rad4-Δ541-Cterm* mutant contains an additional six amino acids (NEAETV) at the end of the C-terminal deletion

APPENDIX C

MODELS AND EXAMPLES OF ANOMALOUS SUBDIFFUSION IN BIOLOGICAL SYSTEMS

The following is only a brief explanation and for a much more thorough discussion on anomalous diffusion; please refer to an excellent recent review of this topic (323).

C.1 CONTINUOUS-TIME RANDOM WALKS

Originally introduced as a stochastic transport model (324) and used to describe motion of charge carriers in amorphous materials (325), the continuous-time random walk (CTRW) model can be considered a generalization of regular Brownian motion. Consider the simplest form of a one-dimensional random walk: a particle makes a jump of step size l, to either the left or the right, after a waiting time t. The CTRW generalization of this description requires both step size and waiting time to be random variables, drawn from separate probability distributions. After each step, a new pair of values for l and t are generated from those same distributions, but independent of the values from the previous step. Naturally, when both distributions are well behaved, i.e. finite variance of

step lengths and mean waiting times, CTRW describes a simple random walk. However, if we assume a power-law form distribution of waiting times,

$$\psi(t) \propto t^{-1-\alpha} \tag{c.1}$$

such that for $0 < \alpha < 1$, the characteristic waiting time $\langle t \rangle$ diverges and the subdiffusive ensemble-averaged MSD takes the form of $\langle x^2(\Delta) \rangle \sim \Delta^{\alpha}$, where Δ is the lag time (326). Furthermore, the lack of a finite characteristic waiting time also leads to what is known as *weak ergodicity breaking* among physicists. In the case of diffusion, a process is considered ergodic if the ensemble average of MSD $\langle x^2(\Delta) \rangle$ and time-averaged MSD $\overline{\delta^2(\Delta)}$ are equivalent in the limit of long measurement times, the latter of which is normally derived from time series collected in biophysical single molecule or single particle tracking experiments. For CTRW specifically (327-329),

$$\langle \overline{\delta^2(\Delta)} \rangle \sim \frac{\Delta}{t^{1-\alpha}} \tag{c.2}$$

The dependence of time-averaged MSD $\langle \overline{\delta^2(\Delta)} \rangle$ on measurement time t leads to the observation of *aging* in the system, such that the time-averaged MSD is smaller if it is measured on a particle that has spent more time undergoing the diffusion process.

Physically, the power-law form of waiting time distribution that results in subdiffusion under the CTRW regime could come from energy traps with exponentially distributed energy wells and their Arrhenius-type escape times (323). CTRW is an attractive model for subdiffusion in biological systems due to the multitude of intermolecular interactions between macromolecules present in the cell that naturally gives rises to energetic traps. Experimentally, CTRW has been shown to accurately model the non-ergodic component of Kv2.1 potassium channels diffusion in two-dimensional plasma membrane, which is caused by transient binding of the potassium channels to the actin cytoskeleton (330). Other experimental evidence lending support to the

CTRW model *in vivo* include observation of subdiffusion of RNA molecules in *E. coli* (331), and that of short time behavior of endogenous lipid granules in living fission yeast cells (332).

C.2 FRACTIONAL BROWNIAN MOTION AND FRACTIONAL LANGEVIN EQUATION

Another major stochastic model of anomalous diffusion is called fractional Brownian motion (FBM), described by Mandelbrot and van Ness (333). FBM is driven by a stationary, fractional Gaussian noise (FGN) with zero mean. Recall that normal diffusion is generated by uncorrelated white noise $\xi(t)$ in Equation 1.8; the FGN is time-difference correlated and the correlation takes on a power-law form

$$\langle \xi_{FGN}(t) \rangle = 0, \langle \xi_{FGN}(t)\xi_{FGN}(t') \rangle = \alpha(\alpha-1)K_\alpha^*|t-t'|^{\alpha-2} \quad \text{(c.3)}$$

When FGN is anti-correlated ($0 < \alpha < 1$), FBM describes subdiffusion.

Based on the generalized Langevin equation (334), and introducing an FGN, as defined in FBM, leads to the fractional Langevin equation (FLE) (335,336). Contrary to FBM, FLE with *correlated* FGN ($1 < \alpha < 2$) leads to FBM-like subdiffusion only in its long time limit (337). Further, in contrast to the characteristic weak ergodicity breaking of CTRW, both FBM and motion governed by FLE have been shown to be ergodic, (i.e., the time averaged MSD converges slowly to the ensemble average), and take on the form of $\sim t^\alpha$ with $0 < \alpha < 1$ (338).

A well-studied FBM/FLE-governed biological system of anomalous diffusion can be found in the subdiffusion of particles in viscoelastic environments, such as the cytoplasm and nucleoplasm of cells, due to effects of molecular crowding (339,340). Evidence from experiments,

as well as simulations of particle subdiffusion in artificially crowded solutions, shows that such a process is most consistent with FBM (341-343). FBM or FLE-governed motion have also been suggested as models for the observed subdiffusion of fluorescently labeled mRNA molecules (344), and chromosomal loci in bacterial cells (345), and transient subdiffusion of telomeres in U2OS nuclei (346,347).

C.3 OBSTRUCTED DIFFUSION

Consider a two-dimensional surface (e.g. a biological membrane), randomly decorated with immobile objects (e.g. anchored proteins) that pose obstacles to free diffusion of particles. This is an intuitive example of obstructed diffusion (OD), one of the simplest models of anomalous diffusion. As obstacle concentration increases, the available space for free diffusion decreases and subdiffusion rises. Mathematical modeling of OD is deeply rooted in percolation theory and diffusion in fractal space (348). Readers are referred to a recent review for more in-depth discussions on the subject (293). At lower obstacle concentrations, subdiffusion is transient before crossing over to normal diffusion. Monte Carlo simulations show that as the obstacle concentration approaches criticality (i.e. the percolation threshold), both crossover time and distance increase, becoming more relevant for observation in biological systems (349). Therefore, experimental observation of obstructed diffusion would appear anomalous over shorter time periods and normal over longer time periods. Like FBM, OD is also ergodic and stationary (293). Simulations of FBM- and OD-based models agreed favorably with experimentally observed molecular crowding-dependent processive phosphorylation of MAP kinase (350,351).

C.4 OTHER MODELS

Finally, subdiffusion can be modeled by assuming a diffusivity that is either time- or position-dependent, namely the scaled Brownian motion (SBM) (352,353) and the heterogeneous diffusion process (HDP) (354,355), respectively. Experimental evidence of the HDP in mammalian cell lines has been established using fluorescence recovery after photobleaching (FRAP) to study freely-diffusing enhanced yellow fluorescent proteins (356). More recently, patch models have been suggested as a family of HDP to explain CTRW-like non-ergodic subdiffusion that results from heterogeneous diffusivity rather than transient trapping (357). Simulations based on the patch model reproduced the observed non-ergodic subdiffusion of receptor on a live cell membrane, where the receptor motion could also be correlated to its structure (358).

C.5 SUBORDINATION

Thus far we have only discussed anomalous diffusion governed solely by a single specific model. However, due to the complexity of biological systems, it is possible that observed behavior is not adequately modeled by any single process. Sometimes, different models are applicable at different time scales. For example, even though endogenous lipid granules in living fission yeast cells undergo short time CTRW subdiffusion, their motion is better described by FBM at longer time scales (332). Other cases require different processes to be combined to form subordination schemes (359,360). Intracellular transport of fluorescently labeled insulin granules was found to be accurately modeled by FBM subordinated to a CTRW (361). Similarly, subordinating the ergodic diffusion on a fractal to a non-ergodic CTRW (362) has been proven appropriate in

modeling subdiffusion of Kv2.1 potassium channels in the plasma membrane (330). These subordinated schemes are likely essential in describing biological systems which are inherently complex and heterogeneous. Future investigations are needed for better understanding of which and to what extent underlying biological processes contribute to distinct mechanisms of subdiffusion.

APPENDIX D

This book chapter is published in B. Eichman (Ed.), *Methods in Enzymology*. Vol. 592. DNA Repair Enzymes: Structure, Biophysics, and Mechanism.

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

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

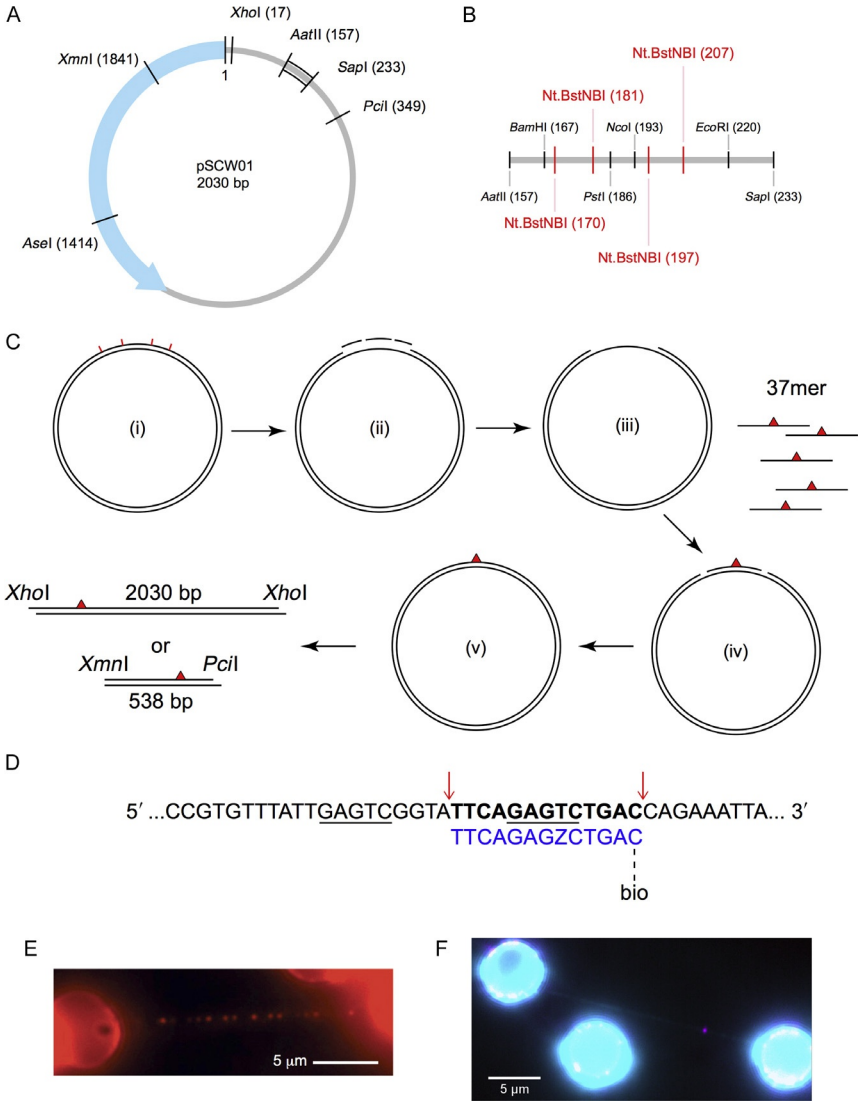

Fig. 1 Design of defined lesion substrate. (A) Map of the pSCW01 plasmid and locations of restriction digest sites. (B) Detailed map of restriction digest and nicking sites for plasmid sequence between the *Aat*II (157) and *Sap*I (233) sites. Nt.BstNBI nicking sites are shown in *red*. (C) Strategy 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

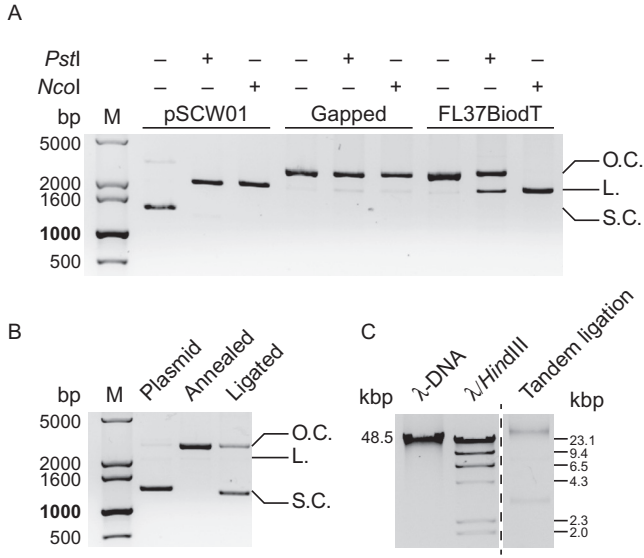

Fig. 2 Diagnostic agarose gels for preparation of defined lesion substrates. (A) 1% agarose gel of diagnostic restriction digests of pSCW01 plasmid, gapped plasmid, and gapped plasmid with FL37BiodT annealed. *Pst*I and *Nco*I, whose restriction sites are within the 37-base gap, do not linearize the 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,

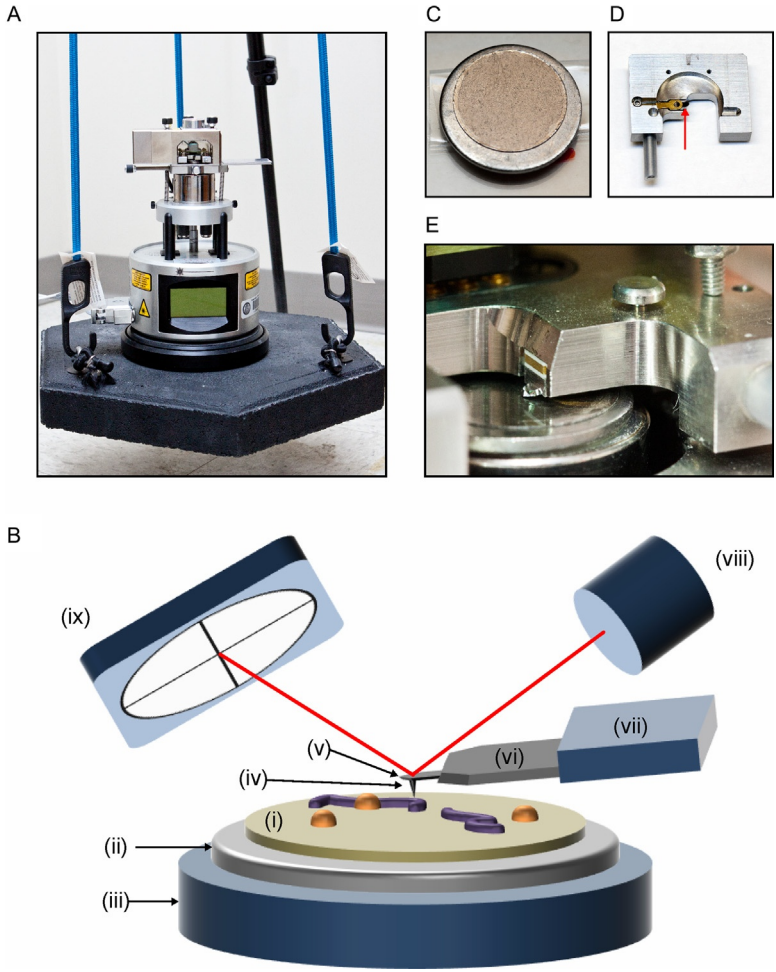

Fig. 3 Atomic force microscopy setup. (A) diMultiMode V atomic force microscope by Veeco. For noise isolation, the AFM is placed on a heavy platform suspended by elastic bungee cords that are secured to a tripod. (B) Schematic of AFM tapping mode in air (not to scale). A protein–DNA sample on mica (i) 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.

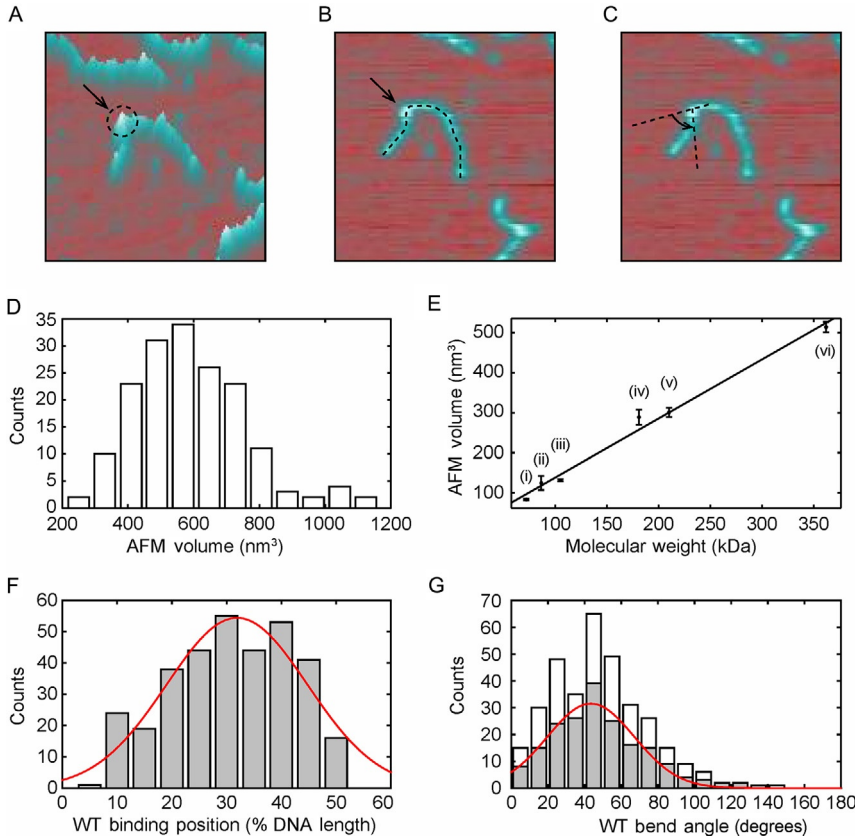

Fig. 4 AFM imaging of protein volume, position, and bend angle. (A)–(C) AFM images of each protein-binding event on DNA can be used to extract information on the protein volume, binding position, and DNA bend angle, respectively. (D) Histogram of UV-DDB K244E mutant volumes on DNA ($n = 171$). (E) Calibration 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, $\sim 5\ \mu M$
- DNA substrate, $\sim 200\ n M$ (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,

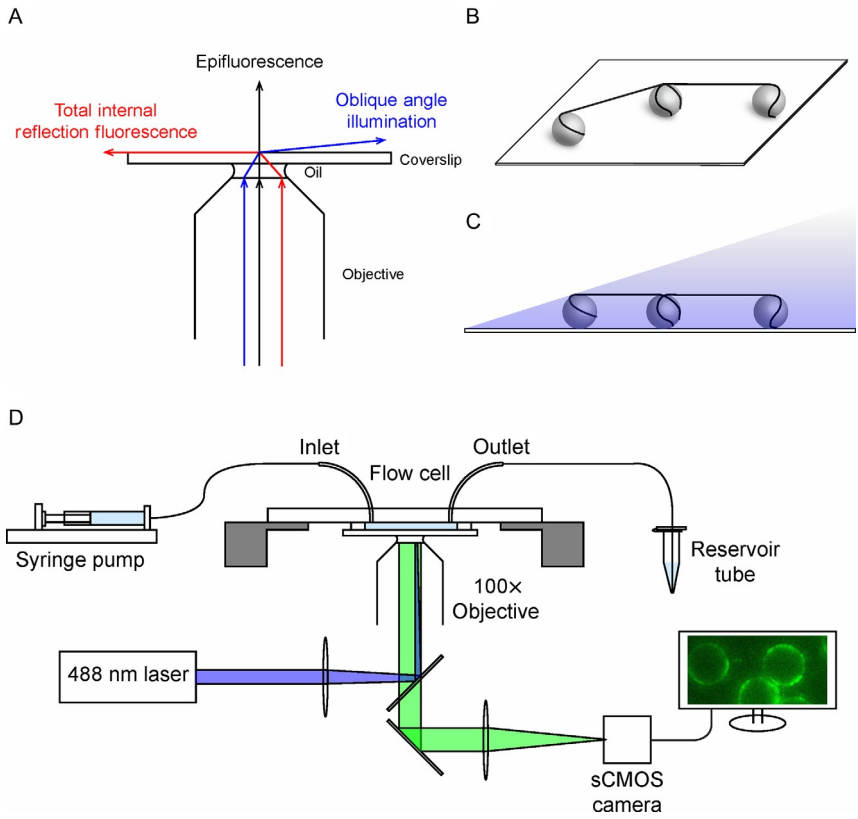

Fig. 5 Schematics of the DNA tightrope assay. (A) Schematic ray diagram of incident laser light paths for epifluorescence (*black*), total internal reflection fluorescence (TIRF) at the critical angle (*red*), and oblique angle illumination (*blue*). (B) Schematic of 5 μm poly-L-lysine-coated microspheres 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

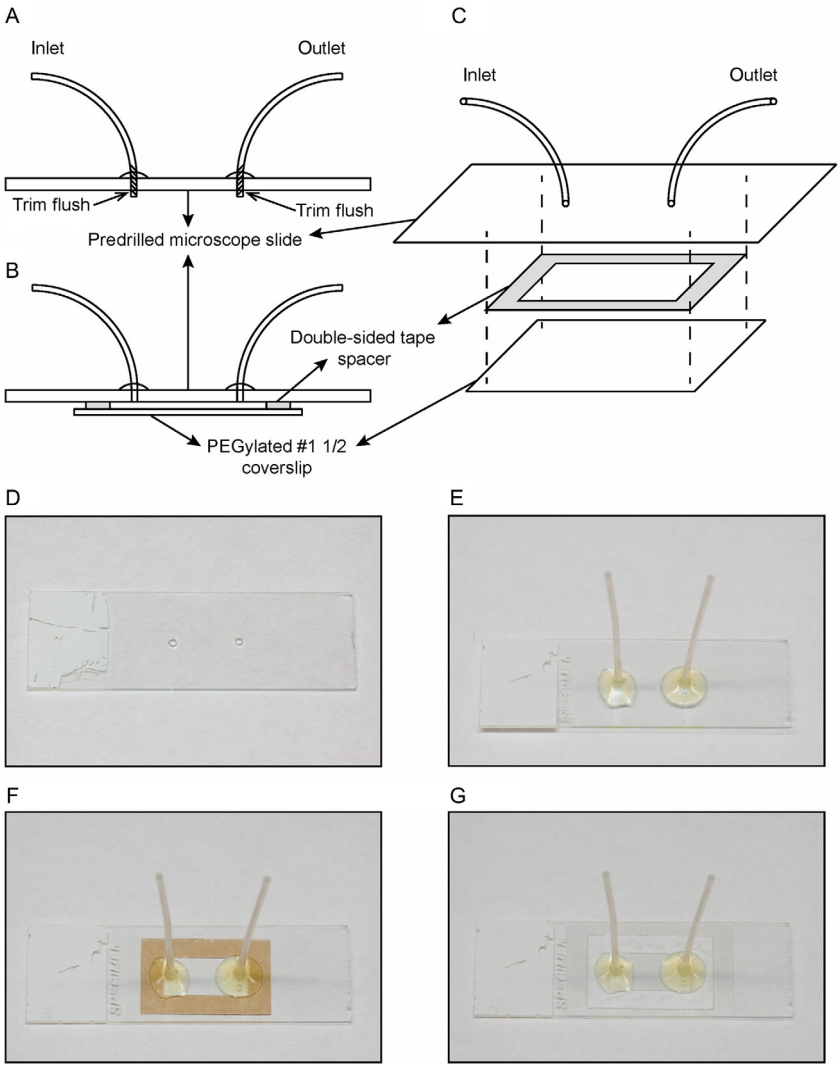

Fig. 6 Assembly of the flow cell. (A) Cross-sectional view of the predrilled microscope slide with inlet and outlet tubing attached. (B) Cross-sectional view of the assembled flow cell, where the double-sided tape spacer is sandwiched between the slide assembly and the PEGylated coverslip. (C) Exploded 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

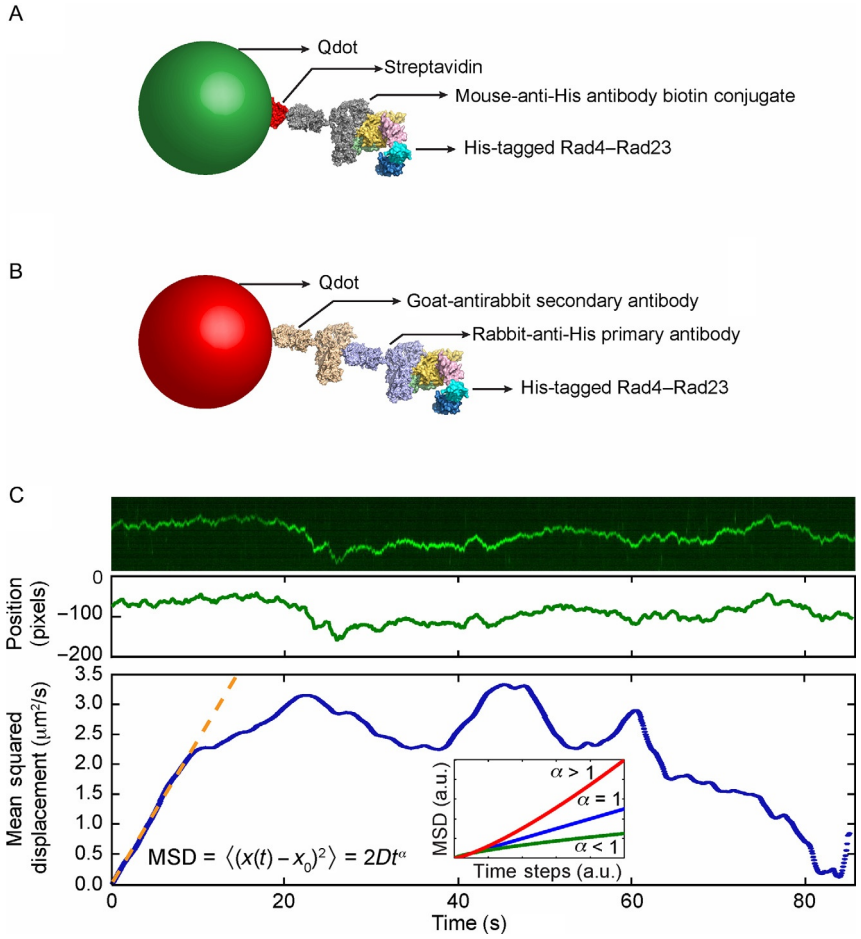

Fig. 7 Qdot conjugation strategies and data analysis. (A) Streptavidin (*red*)-coated quantum dot (*green*) is conjugated to His-tagged Rad4–Rad23 via the biotin-conjugated mouse-anti-His antibody (*gray*). (B) His-tagged Rad4–Rad is labeled by goat-antirabbit secondary antibody (*wheat*)-conjugated 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/nl080316l.

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. Swenberg, J.A., Lu, K., Moeller, B.C., Gao, L., Upton, P.B., Nakamura, J. and Starr, T.B. (2011) Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. *Toxicol Sci*, **120 Suppl 1**, S130-145.

2. Friedberg, E.C. (2011) Nucleotide excision repair of DNA: The very early history. *DNA repair*, **10**, 668-672.

3. Dulbecco, R. (1949) Reactivation of ultra-violet-inactivated bacteriophage by visible light. *Nature*, **163**, 949.

4. Kelner, A. (1949) Effect of Visible Light on the Recovery of Streptomyces Griseus Conidia from Ultra-violet Irradiation Injury. *Proceedings of the National Academy of Sciences of the United States of America*, **35**, 73-79.

5. Friedberg, E.C. (1997) *Correcting the blueprint of life : an historical account of the discovery of DNA repair mechanisms*. Cold Spring Harbor Laboratory Press, Plainview, N.Y.

6. Setlow, R.B. and Carrier, W.L. (1964) The Disappearance of Thymine Dimers from DNA: An Error-Correcting Mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, **51**, 226-231.

7. Boyce, R.P. and Howard-Flanders, P. (1964) Release of Ultraviolet Light-Induced Thymine Dimers from DNA in E. Coli K-12. *Proceedings of the National Academy of Sciences of the United States of America*, **51**, 293-300.

8. Pettijohn, D. and Hanawalt, P. (1964) Evidence for Repair-Replication of Ultraviolet Damaged DNA in Bacteria. *Journal of molecular biology*, **9**, 395-410.

9. Rasmussen, R.E. and Painter, R.B. (1964) Evidence for Repair of Ultra-Violet Damaged Deoxyribonucleic Acid in Cultured Mammalian Cells. *Nature*, **203**, 1360-1362.

10. Cleaver, J.E. (1968) Defective repair replication of DNA in xeroderma pigmentosum. *Nature*, **218**, 652-656.

11. Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071-1078.

12. Mitchell, D.L. and Nairn, R.S. (1989) The biology of the (6-4) photoproduct. *Photochem Photobiol*, **49**, 805-819.

13. Hansson, J., Munn, M., Rupp, W.D., Kahn, R. and Wood, R.D. (1989) Localization of DNA repair synthesis by human cell extracts to a short region at the site of a lesion. *The Journal of biological chemistry*, **264**, 21788-21792.

14. Reardon, J.T., Spielmann, P., Huang, J.C., Sastry, S., Sancar, A. and Hearst, J.E. (1991) Removal of psoralen monoadducts and crosslinks by human cell free extracts. *Nucleic acids research*, **19**, 4623-4629.

15. Huang, J.C., Zamble, D.B., Reardon, J.T., Lippard, S.J. and Sancar, A. (1994) HMG-domain proteins specifically inhibit the repair of the major DNA adduct of the anticancer drug cisplatin by human excision nuclease. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 10394-10398.

16. Hess, M.T., Gunz, D., Luneva, N., Geacintov, N.E. and Naegeli, H. (1997) Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes. *Molecular and cellular biology*, **17**, 7069-7076.

17. Gillet, L.C. and Scharer, O.D. (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chemical reviews*, **106**, 253-276.

18. Mellon, I., Spivak, G. and Hanawalt, P.C. (1987) Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell*, **51**, 241-249.

19. Hanawalt, P.C. and Spivak, G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises. *Nature reviews. Molecular cell biology*, **9**, 958-970.

20. Selby, C.P. and Sancar, A. (1997) Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *The Journal of biological chemistry*, **272**, 1885-1890.

21. van den Boom, V., Citterio, E., Hoogstraten, D., Zotter, A., Egly, J.M., van Cappellen, W.A., Hoeijmakers, J.H., Houtsmuller, A.B. and Vermeulen, W. (2004) DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *The Journal of cell biology*, **166**, 27-36.

22. Beerens, N., Hoeijmakers, J.H., Kanaar, R., Vermeulen, W. and Wyman, C. (2005) The CSB protein actively wraps DNA. *The Journal of biological chemistry*, **280**, 4722-4729.

23. Fousteri, M., Vermeulen, W., van Zeeland, A.A. and Mullenders, L.H. (2006) Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Molecular cell*, **23**, 471-482.

24. Groisman, R., Polanowska, J., Kuraoka, I., Sawada, J., Saijo, M., Drapkin, R., Kisselev, A.F., Tanaka, K. and Nakatani, Y. (2003) The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell*, **113**, 357-367.

25. Groisman, R., Kuraoka, I., Chevallier, O., Gaye, N., Magnaldo, T., Tanaka, K., Kisselev, A.F., Harel-Bellan, A. and Nakatani, Y. (2006) CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes & development*, **20**, 1429-1434.

26. Fei, J. and Chen, J. (2012) KIAA1530 protein is recruited by Cockayne syndrome complementation group protein A (CSA) to participate in transcription-coupled repair (TCR). *The Journal of biological chemistry*, **287**, 35118-35126.

27. Nakazawa, Y., Sasaki, K., Mitsutake, N., Matsuse, M., Shimada, M., Nardo, T., Takahashi, Y., Ohyama, K., Ito, K., Mishima, H. *et al.* (2012) Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair. *Nat Genet*, **44**, 586-592.

28. Schwertman, P., Lagarou, A., Dekkers, D.H., Raams, A., van der Hoek, A.C., Laffeber, C., Hoeijmakers, J.H., Demmers, J.A., Fousteri, M., Vermeulen, W. *et al.* (2012) UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat Genet*, **44**, 598-602.

29. Zhang, X., Horibata, K., Saijo, M., Ishigami, C., Ukai, A., Kanno, S., Tahara, H., Neilan, E.G., Honma, M., Nohmi, T. *et al.* (2012) Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nat Genet*, **44**, 593-597.

30. Tornaletti, S., Reines, D. and Hanawalt, P.C. (1999) Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. *The Journal of biological chemistry*, **274**, 24124-24130.

31. Spivak, G. (2016) Transcription-coupled repair: an update. *Arch Toxicol*, **90**, 2583-2594.

32. Sarker, A.H., Tsutakawa, S.E., Kostek, S., Ng, C., Shin, D.S., Peris, M., Campeau, E., Tainer, J.A., Nogales, E. and Cooper, P.K. (2005) Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome. *Molecular cell*, **20**, 187-198.

33. Wilson, M.D., Harreman, M. and Svejstrup, J.Q. (2013) Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochim Biophys Acta*, **1829**, 151-157.

34. Van Houten, B. and Kong, M. (2016) In Bradshaw, R. A. and Stahl, P. D. (eds.), *Encyclopedia of Cell Biology*. 1 ed. Academic Press, Vol. 1, pp. 435-441.

35. Scharer, O.D. (2013) Nucleotide excision repair in eukaryotes. *Cold Spring Harbor perspectives in biology*, **5**, a012609.

36. Tang, J.Y., Hwang, B.J., Ford, J.M., Hanawalt, P.C. and Chu, G. (2000) Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis. *Molecular cell*, **5**, 737-744.

37. Wakasugi, M., Kawashima, A., Morioka, H., Linn, S., Sancar, A., Mori, T., Nikaido, O. and Matsunaga, T. (2002) DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair. *The Journal of biological chemistry*, **277**, 1637-1640.

38. Fitch, M.E., Nakajima, S., Yasui, A. and Ford, J.M. (2003) In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *The Journal of biological chemistry*, **278**, 46906-46910.

39. Moser, J., Volker, M., Kool, H., Alekseev, S., Vrieling, H., Yasui, A., van Zeeland, A.A. and Mullenders, L.H. (2005) The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions. *DNA repair*, **4**, 571-582.

40. Wang, Q.E., Zhu, Q., Wani, G., Chen, J. and Wani, A.A. (2004) UV radiation-induced XPC translocation within chromatin is mediated by damaged-DNA binding protein, DDB2. *Carcinogenesis*, **25**, 1033-1043.

41. Scrima, A., Konickova, R., Czyzewski, B.K., Kawasaki, Y., Jeffrey, P.D., Groisman, R., Nakatani, Y., Iwai, S., Pavletich, N.P. and Thoma, N.H. (2008) Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell*, **135**, 1213-1223.

42. Wittschieben, B.O., Iwai, S. and Wood, R.D. (2005) DDB1-DDB2 (xeroderma pigmentosum group E) protein complex recognizes a cyclobutane pyrimidine dimer, mismatches, apurinic/apyrimidinic sites, and compound lesions in DNA. *The Journal of biological chemistry*, **280**, 39982-39989.

43. Ghodke, H., Wang, H., Hsieh, C.L., Woldemeskel, S., Watkins, S.C., Rapic-Otrin, V. and Van Houten, B. (2014) Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, E1862-1871.

44. Yeh, J.I., Levine, A.S., Du, S., Chinte, U., Ghodke, H., Wang, H., Shi, H., Hsieh, C.L., Conway, J.F., Van Houten, B. *et al.* (2012) Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E2737-2746.

45. Angers, S., Li, T., Yi, X., MacCoss, M.J., Moon, R.T. and Zheng, N. (2006) Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature*, **443**, 590-593.

46. Fischer, E.S., Scrima, A., Bohm, K., Matsumoto, S., Lingaraju, G.M., Faty, M., Yasuda, T., Cavadini, S., Wakasugi, M., Hanaoka, F. *et al.* (2011) The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell*, **147**, 1024-1039.

47. Scrima, A., Fischer, E.S., Lingaraju, G.M., Bohm, K., Cavadini, S. and Thoma, N.H. (2011) Detecting UV-lesions in the genome: The modular CRL4 ubiquitin ligase does it best! *FEBS Lett*, **585**, 2818-2825.

48. Wang, H., Zhai, L., Xu, J., Joo, H.Y., Jackson, S., Erdjument-Bromage, H., Tempst, P., Xiong, Y. and Zhang, Y. (2006) Histone H3 and H4 ubiquitylation by the CUL4-DDB-ROC1 ubiquitin ligase facilitates cellular response to DNA damage. *Molecular cell*, **22**, 383-394.

49. Kapetanaki, M.G., Guerrero-Santoro, J., Bisi, D.C., Hsieh, C.L., Rapic-Otrin, V. and Levine, A.S. (2006) The DDB1-CUL4ADDB2 ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 2588-2593.

50. Lan, L., Nakajima, S., Kapetanaki, M.G., Hsieh, C.L., Fagerburg, M., Thickman, K., Rodriguez-Collazo, P., Leuba, S.H., Levine, A.S. and Rapic-Otrin, V. (2012) Monoubiquitinated histone H2A destabilizes photolesion-containing nucleosomes with concomitant release of UV-damaged DNA-binding protein E3 ligase. *The Journal of biological chemistry*, **287**, 12036-12049.

51. Sugasawa, K., Okuda, Y., Saijo, M., Nishi, R., Matsuda, N., Chu, G., Mori, T., Iwai, S., Tanaka, K., Tanaka, K. *et al.* (2005) UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell*, **121**, 387-400.

52. Sugasawa, K. (2006) UV-induced ubiquitylation of XPC complex, the UV-DDB-ubiquitin ligase complex, and DNA repair. *Journal of molecular histology*, **37**, 189-202.

53. Wang, Q.E., Zhu, Q., Wani, G., El-Mahdy, M.A., Li, J. and Wani, A.A. (2005) DNA repair factor XPC is modified by SUMO-1 and ubiquitin following UV irradiation. *Nucleic acids research*, **33**, 4023-4034.

54. Wang, Q.E., Praetorius-Ibba, M., Zhu, Q., El-Mahdy, M.A., Wani, G., Zhao, Q., Qin, S., Patnaik, S. and Wani, A.A. (2007) Ubiquitylation-independent degradation of Xeroderma pigmentosum group C protein is required for efficient nucleotide excision repair. *Nucleic acids research*, **35**, 5338-5350.

55. Matsumoto, S., Fischer, E.S., Yasuda, T., Dohmae, N., Iwai, S., Mori, T., Nishi, R., Yoshino, K., Sakai, W., Hanaoka, F. *et al.* (2015) Functional regulation of the DNA damage-recognition factor DDB2 by ubiquitination and interaction with xeroderma pigmentosum group C protein. *Nucleic acids research*, **43**, 1700-1713.

56. Zhang, L., Lubin, A., Chen, H., Sun, Z. and Gong, F. (2012) The deubiquitinating protein USP24 interacts with DDB2 and regulates DDB2 stability. *Cell cycle*, **11**, 4378-4384.

57. Pines, A., Vrouwe, M.G., Marteijn, J.A., Typas, D., Luijsterburg, M.S., Cansoy, M., Hensbergen, P., Deelder, A., de Groot, A., Matsumoto, S. *et al.* (2012) PARP1 promotes nucleotide excision repair through DDB2 stabilization and recruitment of ALC1. *The Journal of cell biology*, **199**, 235-249.

58. Robu, M., Shah, R.G., Petitclerc, N., Brind'Amour, J., Kandan-Kulangara, F. and Shah, G.M. (2013) Role of poly(ADP-ribose) polymerase-1 in the removal of UV-induced DNA lesions by nucleotide excision repair. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 1658-1663.

59. van der Spek, P.J., Smit, E.M., Beverloo, H.B., Sugasawa, K., Masutani, C., Hanaoka, F., Hoeijmakers, J.H. and Hagemeijer, A. (1994) Chromosomal localization of three repair genes: the xeroderma pigmentosum group C gene and two human homologs of yeast RAD23. *Genomics*, **23**, 651-658.

60. Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P.J., Bootsma, D. *et al.* (1994) Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *The EMBO journal*, **13**, 1831-1843.

61. Sugasawa, K., Masutani, C., Uchida, A., Maekawa, T., van der Spek, P.J., Bootsma, D., Hoeijmakers, J.H. and Hanaoka, F. (1996) HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro. *Molecular and cellular biology*, **16**, 4852-4861.

62. Sugasawa, K., Ng, J.M., Masutani, C., Maekawa, T., Uchida, A., van der Spek, P.J., Eker, A.P., Rademakers, S., Visser, C., Aboussekhra, A. *et al.* (1997) Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity. *Molecular and cellular biology*, **17**, 6924-6931.

63. Ng, J.M., Vermeulen, W., van der Horst, G.T., Bergink, S., Sugasawa, K., Vrieling, H. and Hoeijmakers, J.H. (2003) A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes & development*, **17**, 1630-1645.

64. Okuda, Y., Nishi, R., Ng, J.M., Vermeulen, W., van der Horst, G.T., Mori, T., Hoeijmakers, J.H., Hanaoka, F. and Sugasawa, K. (2004) Relative levels of the two mammalian Rad23 homologs determine composition and stability of the xeroderma pigmentosum group C protein complex. *DNA repair*, **3**, 1285-1295.

65. Yokoi, M. and Hanaoka, F. (2017) Two mammalian homologs of yeast Rad23, HR23A and HR23B, as multifunctional proteins. *Gene*, **597**, 1-9.

66. Araki, M., Masutani, C., Takemura, M., Uchida, A., Sugasawa, K., Kondoh, J., Ohkuma, Y. and Hanaoka, F. (2001) Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. *The Journal of biological chemistry*, **276**, 18665-18672.

67. Nishi, R., Okuda, Y., Watanabe, E., Mori, T., Iwai, S., Masutani, C., Sugasawa, K. and Hanaoka, F. (2005) Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Molecular and cellular biology*, **25**, 5664-5674.

68. Popescu, A., Miron, S., Blouquit, Y., Duchambon, P., Christova, P. and Craescu, C.T. (2003) Xeroderma pigmentosum group C protein possesses a high affinity binding site to human centrin 2 and calmodulin. *The Journal of biological chemistry*, **278**, 40252-40261.

69. Nishi, R., Sakai, W., Tone, D., Hanaoka, F. and Sugasawa, K. (2013) Structure-function analysis of the EF-hand protein centrin-2 for its intracellular localization and nucleotide excision repair. *Nucleic acids research*, **41**, 6917-6929.

70. Miron, S., Duchambon, P., Blouquit, Y., Durand, D. and Craescu, C.T. (2008) The carboxy-terminal domain of xeroderma pigmentosum complementation group C protein, involved in TFIIH and centrin binding, is highly disordered. *Biochemistry*, **47**, 1403-1413.

71. Sugasawa, K., Ng, J.M., Masutani, C., Iwai, S., van der Spek, P.J., Eker, A.P., Hanaoka, F., Bootsma, D. and Hoeijmakers, J.H. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Molecular cell*, **2**, 223-232.

72. Bergink, S., Toussaint, W., Luijsterburg, M.S., Dinant, C., Alekseev, S., Hoeijmakers, J.H., Dantuma, N.P., Houtsmuller, A.B. and Vermeulen, W. (2012) Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex. *The Journal of cell biology*, **196**, 681-688.

73. Volker, M., Mone, M.J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J.H., van Driel, R., van Zeeland, A.A. and Mullenders, L.H. (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Molecular cell*, **8**, 213-224.

74. Riedl, T., Hanaoka, F. and Egly, J.M. (2003) The comings and goings of nucleotide excision repair factors on damaged DNA. *The EMBO journal*, **22**, 5293-5303.

75. Tapias, A., Auriol, J., Forget, D., Enzlin, J.H., Scharer, O.D., Coin, F., Coulombe, B. and Egly, J.M. (2004) Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *The Journal of biological chemistry*, **279**, 19074-19083.

76. Kusumoto, R., Masutani, C., Sugasawa, K., Iwai, S., Araki, M., Uchida, A., Mizukoshi, T. and Hanaoka, F. (2001) Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro. *Mutation research*, **485**, 219-227.

77. Trego, K.S. and Turchi, J.J. (2006) Pre-steady-state binding of damaged DNA by XPC-hHR23B reveals a kinetic mechanism for damage discrimination. *Biochemistry*, **45**, 1961-1969.

78. Hey, T., Lipps, G., Sugasawa, K., Iwai, S., Hanaoka, F. and Krauss, G. (2002) The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay. *Biochemistry*, **41**, 6583-6587.

79. Mocquet, V., Kropachev, K., Kolbanovskiy, M., Kolbanovskiy, A., Tapias, A., Cai, Y., Broyde, S., Geacintov, N.E. and Egly, J.M. (2007) The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. *The EMBO journal*, **26**, 2923-2932.

80. Batty, D., Rapic'-Otrin, V., Levine, A.S. and Wood, R.D. (2000) Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *Journal of molecular biology*, **300**, 275-290.

81. Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S. and Hanaoka, F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes & development*, **15**, 507-521.

82. Sugasawa, K., Shimizu, Y., Iwai, S. and Hanaoka, F. (2002) A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA repair*, **1**, 95-107.

83. Roche, Y., Zhang, D., Segers-Nolten, G.M., Vermeulen, W., Wyman, C., Sugasawa, K., Hoeijmakers, J. and Otto, C. (2008) Fluorescence correlation spectroscopy of the binding of nucleotide excision repair protein XPC-hHr23B with DNA substrates. *Journal of fluorescence*, **18**, 987-995.

84. Brown, K.L., Roginskaya, M., Zou, Y., Altamirano, A., Basu, A.K. and Stone, M.P. (2010) 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 research*, **38**, 428-440.

85. Maillard, O., Camenisch, U., Clement, F.C., Blagoev, K.B. and Naegeli, H. (2007) DNA repair triggered by sensors of helical dynamics. *Trends Biochem Sci*, **32**, 494-499.

86. Legerski, R. and Peterson, C. (1992) Expression cloning of a human DNA repair gene involved in xeroderma pigmentosum group C. *Nature*, **359**, 70-73.

87. Henning, K.A., Peterson, C., Legerski, R. and Friedberg, E.C. (1994) Cloning the Drosophila homolog of the xeroderma pigmentosum complementation group C gene reveals homology between the predicted human and Drosophila polypeptides and that encoded by the yeast RAD4 gene. *Nucleic acids research*, **22**, 257-261.

88. Min, J.H. and Pavletich, N.P. (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature*, **449**, 570-575.

89. Kim, J.K., Patel, D. and Choi, B.S. (1995) Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity. *Photochem Photobiol*, **62**, 44-50.

90. Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K. *et al.* (2015) The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell*, **162**, 425-440.

91. Maltseva, E.A., Rechkunova, N.I., Sukhanova, M.V. and Lavrik, O.I. (2015) Poly(ADP-ribose) Polymerase 1 Modulates Interaction of the Nucleotide Excision Repair Factor XPC-RAD23B with DNA via Poly(ADP-ribosyl)ation. *The Journal of biological chemistry*, **290**, 21811-21820.

92. Poulsen, S.L., Hansen, R.K., Wagner, S.A., van Cuijk, L., van Belle, G.J., Streicher, W., Wikstrom, M., Choudhary, C., Houtsmuller, A.B., Marteijn, J.A. *et al.* (2013) RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. *The Journal of cell biology*, **201**, 797-807.

93. van Cuijk, L., van Belle, G.J., Turkyilmaz, Y., Poulsen, S.L., Janssens, R.C., Theil, A.F., Sabatella, M., Lans, H., Mailand, N., Houtsmuller, A.B. *et al.* (2015) SUMO and ubiquitin-dependent XPC exchange drives nucleotide excision repair. *Nature communications*, **6**, 7499.

94. Akita, M., Tak, Y.S., Shimura, T., Matsumoto, S., Okuda-Shimizu, Y., Shimizu, Y., Nishi, R., Saitoh, H., Iwai, S., Mori, T. *et al.* (2015) SUMOylation of xeroderma pigmentosum group C protein regulates DNA damage recognition during nucleotide excision repair. *Sci Rep*, **5**, 10984.

95. Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y. and Hanaoka, F. (2000) The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. *The Journal of biological chemistry*, **275**, 9870-9875.

96. Uchida, A., Sugasawa, K., Masutani, C., Dohmae, N., Araki, M., Yokoi, M., Ohkuma, Y. and Hanaoka, F. (2002) The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH. *DNA repair*, **1**, 449-461.

97. Araujo, S.J., Nigg, E.A. and Wood, R.D. (2001) Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. *Molecular and cellular biology*, **21**, 2281-2291.

98. Compe, E. and Egly, J.M. (2016) Nucleotide Excision Repair and Transcriptional Regulation: TFIIH and Beyond. *Annu Rev Biochem*, **85**, 265-290.

99. Bernardes de Jesus, B.M., Bjoras, M., Coin, F. and Egly, J.M. (2008) Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. *Molecular and cellular biology*, **28**, 7225-7235.

100. Okuda, M., Kinoshita, M., Kakumu, E., Sugasawa, K. and Nishimura, Y. (2015) Structural Insight into the Mechanism of TFIIH Recognition by the Acidic String of the Nucleotide Excision Repair Factor XPC. *Structure*, **23**, 1827-1837.

101. Araujo, S.J., Tirode, F., Coin, F., Pospiech, H., Syvaoja, J.E., Stucki, M., Hubscher, U., Egly, J.M. and Wood, R.D. (2000) Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes & development*, **14**, 349-359.

102. Coin, F., Oksenych, V., Mocquet, V., Groh, S., Blattner, C. and Egly, J.M. (2008) Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Molecular cell*, **31**, 9-20.

103. Coin, F., Oksenych, V. and Egly, J.M. (2007) Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Molecular cell*, **26**, 245-256.

104. Oksenych, V., Bernardes de Jesus, B., Zhovmer, A., Egly, J.M. and Coin, F. (2009) Molecular insights into the recruitment of TFIIH to sites of DNA damage. *The EMBO journal*, **28**, 2971-2980.

105. Winkler, G.S., Araujo, S.J., Fiedler, U., Vermeulen, W., Coin, F., Egly, J.M., Hoeijmakers, J.H., Wood, R.D., Timmers, H.T. and Weeda, G. (2000) TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *The Journal of biological chemistry*, **275**, 4258-4266.

106. Sugasawa, K., Akagi, J., Nishi, R., Iwai, S. and Hanaoka, F. (2009) Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. *Molecular cell*, **36**, 642-653.

107. Fan, L., Fuss, J.O., Cheng, Q.J., Arvai, A.S., Hammel, M., Roberts, V.A., Cooper, P.K. and Tainer, J.A. (2008) XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell*, **133**, 789-800.

108. Liu, H., Rudolf, J., Johnson, K.A., McMahon, S.A., Oke, M., Carter, L., McRobbie, A.M., Brown, S.E., Naismith, J.H. and White, M.F. (2008) Structure of the DNA repair helicase XPD. *Cell*, **133**, 801-812.

109. Wolski, S.C., Kuper, J., Hanzelmann, P., Truglio, J.J., Croteau, D.L., Van Houten, B. and Kisker, C. (2008) Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS biology*, **6**, e149.

110. Kuper, J., Wolski, S.C., Michels, G. and Kisker, C. (2012) Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *The EMBO journal*, **31**, 494-502.

111. Mathieu, N., Kaczmarek, N. and Naegeli, H. (2010) Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 17545-17550.

112. Abdulrahman, W., Iltis, I., Radu, L., Braun, C., Maglott-Roth, A., Giraudon, C., Egly, J.M. and Poterszman, A. (2013) ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, E633-642.

113. Wakasugi, M. and Sancar, A. (1999) Order of assembly of human DNA repair excision nuclease. *The Journal of biological chemistry*, **274**, 18759-18768.

114. Sugitani, N., Sivley, R.M., Perry, K.E., Capra, J.A. and Chazin, W.J. (2016) XPA: A key scaffold for human nucleotide excision repair. *DNA repair*, **44**, 123-135.

115. Wakasugi, M., Kasashima, H., Fukase, Y., Imura, M., Imai, R., Yamada, S., Cleaver, J.E. and Matsunaga, T. (2009) Physical and functional interaction between DDB and XPA in nucleotide excision repair. *Nucleic acids research*, **37**, 516-525.

116. You, J.S., Wang, M. and Lee, S.H. (2003) Biochemical analysis of the damage recognition process in nucleotide excision repair. *The Journal of biological chemistry*, **278**, 7476-7485.

117. Bunick, C.G., Miller, M.R., Fuller, B.E., Fanning, E. and Chazin, W.J. (2006) Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein. *Biochemistry*, **45**, 14965-14979.

118. Park, C.H., Mu, D., Reardon, J.T. and Sancar, A. (1995) The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIE transcription factor. *The Journal of biological chemistry*, **270**, 4896-4902.

119. Nocentini, S., Coin, F., Saijo, M., Tanaka, K. and Egly, J.M. (1997) DNA damage recognition by XPA protein promotes efficient recruitment of transcription factor II H. *The Journal of biological chemistry*, **272**, 22991-22994.

120. Ziani, S., Nagy, Z., Alekseev, S., Soutoglou, E., Egly, J.M. and Coin, F. (2014) Sequential and ordered assembly of a large DNA repair complex on undamaged chromatin. *The Journal of cell biology*, **206**, 589-598.

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

122. Mer, G., Bochkarev, A., Gupta, R., Bochkareva, E., Frappier, L., Ingles, C.J., Edwards, A.M. and Chazin, W.J. (2000) Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell*, **103**, 449-456.

123. Buchko, G.W., Daughdrill, G.W., de Lorimier, R., Rao, B.K., Isern, N.G., Lingbeck, J.M., Taylor, J.S., Wold, M.S., Gochin, M., Spicer, L.D. *et al.* (1999) 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.

124. Daughdrill, G.W., Buchko, G.W., Botuyan, M.V., Arrowsmith, C., Wold, M.S., Kennedy, M.A. and Lowry, D.F. (2003) 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 research*, **31**, 4176-4183.

125. Saijo, M., Takedachi, A. and Tanaka, K. (2011) Nucleotide excision repair by mutant xeroderma pigmentosum group A (XPA) proteins with deficiency in interaction with RPA. *The Journal of biological chemistry*, **286**, 5476-5483.

126. Li, L., Elledge, S.J., Peterson, C.A., Bales, E.S. and Legerski, R.J. (1994) 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.

127. Li, L., Peterson, C.A., Lu, X. and Legerski, R.J. (1995) Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Molecular and cellular biology*, **15**, 1993-1998.

128. Tsodikov, O.V., Ivanov, D., Orelli, B., Staresincic, L., Shoshani, I., Oberman, R., Scharer, O.D., Wagner, G. and Ellenberger, T. (2007) Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA. *The EMBO journal*, **26**, 4768-4776.

129. Gilljam, K.M., Muller, R., Liabakk, N.B. and Otterlei, M. (2012) Nucleotide excision repair is associated with the replisome and its efficiency depends on a direct interaction between XPA and PCNA. *PloS one*, **7**, e49199.

130. Jones, C.J. and Wood, R.D. (1993) Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochemistry*, **32**, 12096-12104.

131. Liu, Y., Liu, Y., Yang, Z., Utzat, C., Wang, G., Basu, A.K. and Zou, Y. (2005) Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage. *Biochemistry*, **44**, 7361-7368.

132. Yang, Z., Roginskaya, M., Colis, L.C., Basu, A.K., Shell, S.M., Liu, Y., Musich, P.R., Harris, C.M., Harris, T.M. and Zou, Y. (2006) 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.

133. Missura, M., Buterin, T., Hindges, R., Hubscher, U., Kasparkova, J., Brabec, V. and Naegeli, H. (2001) Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair. *The EMBO journal*, **20**, 3554-3564.

134. Hey, T., Lipps, G. and Krauss, G. (2001) Binding of XPA and RPA to damaged DNA investigated by fluorescence anisotropy. *Biochemistry*, **40**, 2901-2910.

135. Camenisch, U., Dip, R., Vitanescu, M. and Naegeli, H. (2007) Xeroderma pigmentosum complementation group A protein is driven to nucleotide excision repair sites by the electrostatic potential of distorted DNA. *DNA repair*, **6**, 1819-1828.

136. Camenisch, U., Dip, R., Schumacher, S.B., Schuler, B. and Naegeli, H. (2006) Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nature structural & molecular biology*, **13**, 278-284.

137. Li, C.L., Golebiowski, F.M., Onishi, Y., Samara, N.L., Sugasawa, K. and Yang, W. (2015) Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Molecular cell*, **59**, 1025-1034.

138. Krasikova, Y.S., Rechkunova, N.I., Maltseva, E.A., Petruseva, I.O. and Lavrik, O.I. (2010) Localization of xeroderma pigmentosum group A protein and replication protein A on damaged DNA in nucleotide excision repair. *Nucleic acids research*, **38**, 8083-8094.

139. de Laat, W.L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N.G. and Hoeijmakers, J.H. (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes & development*, **12**, 2598-2609.

140. Dunand-Sauthier, I., Hohl, M., Thorel, F., Jaquier-Gubler, P., Clarkson, S.G. and Scharer, O.D. (2005) The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. *The Journal of biological chemistry*, **280**, 7030-7037.

141. Zotter, A., Luijsterburg, M.S., Warmerdam, D.O., Ibrahim, S., Nigg, A., van Cappellen, W.A., Hoeijmakers, J.H., van Driel, R., Vermeulen, W. and Houtsmuller, A.B. (2006) Recruitment of the nucleotide excision repair endonuclease XPG to sites of UV-induced dna damage depends on functional TFIIH. *Molecular and cellular biology*, **26**, 8868-8879.

142. Ito, S., Kuraoka, I., Chymkowitch, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J.M. and Tanaka, K. (2007) XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: implications for Cockayne syndrome in XP-G/CS patients. *Molecular cell*, **26**, 231-243.

143. Biggerstaff, M., Szymkowski, D.E. and Wood, R.D. (1993) Co-correction of the ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects in vitro. *The EMBO journal*, **12**, 3685-3692.

144. van Vuuren, A.J., Appeldoorn, E., Odijk, H., Yasui, A., Jaspers, N.G., Bootsma, D. and Hoeijmakers, J.H. (1993) Evidence for a repair enzyme complex involving ERCC1 and complementing activities of ERCC4, ERCC11 and xeroderma pigmentosum group F. *The EMBO journal*, **12**, 3693-3701.

145. Sijbers, A.M., de Laat, W.L., Ariza, R.R., Biggerstaff, M., Wei, Y.F., Moggs, J.G., Carter, K.C., Shell, B.K., Evans, E., de Jong, M.C. *et al.* (1996) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell*, **86**, 811-822.

146. Bardwell, A.J., Bardwell, L., Tomkinson, A.E. and Friedberg, E.C. (1994) Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. *Science*, **265**, 2082-2085.

147. de Laat, W.L., Appeldoorn, E., Jaspers, N.G. and Hoeijmakers, J.H. (1998) DNA structural elements required for ERCC1-XPF endonuclease activity. *The Journal of biological chemistry*, **273**, 7835-7842.

148. Bessho, T., Sancar, A., Thompson, L.H. and Thelen, M.P. (1997) Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex. *The Journal of biological chemistry*, **272**, 3833-3837.

149. Brookman, K.W., Lamerdin, J.E., Thelen, M.P., Hwang, M., Reardon, J.T., Sancar, A., Zhou, Z.Q., Walter, C.A., Parris, C.N. and Thompson, L.H. (1996) ERCC4 (XPF) encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Molecular and cellular biology*, **16**, 6553-6562.

150. Hohl, M., Thorel, F., Clarkson, S.G. and Scharer, O.D. (2003) Structural determinants for substrate binding and catalysis by the structure-specific endonuclease XPG. *The Journal of biological chemistry*, **278**, 19500-19508.

151. Staresincic, L., Fagbemi, A.F., Enzlin, J.H., Gourdin, A.M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S.G., Vermeulen, W. and Scharer, O.D. (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *The EMBO journal*, **28**, 1111-1120.

152. Kemp, M.G., Reardon, J.T., Lindsey-Boltz, L.A. and Sancar, A. (2012) Mechanism of release and fate of excised oligonucleotides during nucleotide excision repair. *The Journal of biological chemistry*, **287**, 22889-22899.

153. Fagbemi, A.F., Orelli, B. and Scharer, O.D. (2011) Regulation of endonuclease activity in human nucleotide excision repair. *DNA repair*, **10**, 722-729.

154. Shivji, M.K., Podust, V.N., Hubscher, U. and Wood, R.D. (1995) Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry*, **34**, 5011-5017.

155. Ogi, T. and Lehmann, A.R. (2006) The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair. *Nature cell biology*, **8**, 640-642.

156. Ogi, T., Limsirichaikul, S., Overmeer, R.M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N.G. *et al.* (2010) Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Molecular cell*, **37**, 714-727.

157. Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L.H. and Fousteri, M.I. (2007) Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Molecular cell*, **27**, 311-323.

158. Kraemer, K.H. and Rünger, T.M. (2008) In Wolff, K., Goldsmith, L. A., Katz, S. I., Gilchrest, B. A., Paller, A. S. and Leffell, D. J. (eds.), *Fitzpatrick's Dermatology in General Medicine*. 7th ed. McGraw-Hill, New York, pp. 977-986.

159. Kleijer, W.J., Laugel, V., Berneburg, M., Nardo, T., Fawcett, H., Gratchev, A., Jaspers, N.G., Sarasin, A., Stefanini, M. and Lehmann, A.R. (2008) Incidence of DNA repair deficiency disorders in western Europe: Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *DNA repair*, **7**, 744-750.

160. Hirai, Y., Kodama, Y., Moriwaki, S., Noda, A., Cullings, H.M., Macphee, D.G., Kodama, K., Mabuchi, K., Kraemer, K.H., Land, C.E. *et al.* (2006) Heterozygous individuals bearing a founder mutation in the XPA DNA repair gene comprise nearly 1% of the Japanese population. *Mutation research*, **601**, 171-178.

161. DiGiovanna, J.J. and Kraemer, K.H. (2012) Shining a light on xeroderma pigmentosum. *The Journal of investigative dermatology*, **132**, 785-796.

162. Tanaka, K., Miura, N., Satokata, I., Miyamoto, I., Yoshida, M.C., Satoh, Y., Kondo, S., Yasui, A., Okayama, H. and Okada, Y. (1990) Analysis of a human DNA excision repair gene involved in group A xeroderma pigmentosum and containing a zinc-finger domain. *Nature*, **348**, 73-76.

163. Weeda, G., van Ham, R.C., Vermeulen, W., Bootsma, D., van der Eb, A.J. and Hoeijmakers, J.H. (1990) A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. *Cell*, **62**, 777-791.

164. Weeda, G., van Ham, R.C., Masurel, R., Westerveld, A., Odijk, H., de Wit, J., Bootsma, D., van der Eb, A.J. and Hoeijmakers, J.H. (1990) Molecular cloning and biological characterization of the human excision repair gene ERCC-3. *Molecular and cellular biology*, **10**, 2570-2581.

165. Flejter, W.L., McDaniel, L.D., Askari, M., Friedberg, E.C. and Schultz, R.A. (1992) Characterization of a complex chromosomal rearrangement maps the locus for in vitro complementation of xeroderma pigmentosum group D to human chromosome band 19q13. *Genes Chromosomes Cancer*, **5**, 335-342.

166. Flejter, W.L., McDaniel, L.D., Johns, D., Friedberg, E.C. and Schultz, R.A. (1992) Correction of xeroderma pigmentosum complementation group D mutant cell phenotypes by chromosome and gene transfer: involvement of the human ERCC2 DNA repair gene. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 261-265.

167. Dualan, R., Brody, T., Keeney, S., Nichols, A.F., Admon, A. and Linn, S. (1995) 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**, 62-69.

168. Mudgett, J.S. and MacInnes, M.A. (1990) Isolation of the functional human excision repair gene ERCC5 by intercosmid recombination. *Genomics*, **8**, 623-633.

169. Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K. and Hanaoka, F. (1999) The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature*, **399**, 700-704.

170. Johnson, R.E., Kondratick, C.M., Prakash, S. and Prakash, L. (1999) hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science*, **285**, 263-265.

171. Bradford, P.T., Goldstein, A.M., Tamura, D., Khan, S.G., Ueda, T., Boyle, J., Oh, K.S., Imoto, K., Inui, H., Moriwaki, S. *et al.* (2011) Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair. *J Med Genet*, **48**, 168-176.

172. Kraemer, K.H., Lee, M.M., Andrews, A.D. and Lambert, W.C. (1994) The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. *Arch Dermatol*, **130**, 1018-1021.

173. Kraemer, K.H., Patronas, N.J., Schiffmann, R., Brooks, B.P., Tamura, D. and DiGiovanna, J.J. (2007) Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship. *Neuroscience*, **145**, 1388-1396.

174. Sethi, M., Lehmann, A.R., Fawcett, H., Stefanini, M., Jaspers, N., Mullard, K., Turner, S., Robson, A., McGibbon, D., Sarkany, R. *et al.* (2013) Patients with xeroderma pigmentosum complementation groups C, E and V do not have abnormal sunburn reactions. *Br J Dermatol*, **169**, 1279-1287.

175. Fassihi, H., Sethi, M., Fawcett, H., Wing, J., Chandler, N., Mohammed, S., Craythorne, E., Morley, A.M., Lim, R., Turner, S. *et al.* (2016) Deep phenotyping of 89 xeroderma pigmentosum patients reveals unexpected heterogeneity dependent on the precise molecular defect. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E1236-1245.

176. Li, L., Bales, E.S., Peterson, C.A. and Legerski, R.J. (1993) Characterization of molecular defects in xeroderma pigmentosum group C. *Nat Genet*, **5**, 413-417.

177. Chavanne, F., Broughton, B.C., Pietra, D., Nardo, T., Browitt, A., Lehmann, A.R. and Stefanini, M. (2000) Mutations in the XPC gene in families with xeroderma pigmentosum and consequences at the cell, protein, and transcript levels. *Cancer Res*, **60**, 1974-1982.

178. Jacobelli, S., Soufir, N., Lacapere, J.J., Regnier, S., Bourillon, A., Grandchamp, B., Hetet, G., Pham, D., Palangie, A., Avril, M.F. *et al.* (2008) Xeroderma pigmentosum group C in a French Caucasian patient with multiple melanoma and unusual long-term survival. *Br J Dermatol*, **159**, 968-973.

179. Meneses, M., Chavez-Bourgeois, M., Badenas, C., Villablanca, S., Aguilera, P., Bennassar, A., Alos, L., Puig, S., Malvehy, J. and Carrera, C. (2015) Atypical Clinical Presentation of Xeroderma Pigmentosum in a Patient Harboring a Novel Missense Mutation in the XPC Gene: The Importance of Clinical Suspicion. *Dermatology*, **231**, 217-221.

180. Nance, M.A. and Berry, S.A. (1992) Cockayne syndrome: review of 140 cases. *Am J Med Genet*, **42**, 68-84.

181. Cleaver, J.E., Lam, E.T. and Revet, I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat Rev Genet*, **10**, 756-768.

182. Nardo, T., Oneda, R., Spivak, G., Vaz, B., Mortier, L., Thomas, P., Orioli, D., Laugel, V., Stary, A., Hanawalt, P.C. *et al.* (2009) A UV-sensitive syndrome patient with a specific CSA mutation reveals separable roles for CSA in response to UV and oxidative DNA damage. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 6209-6214.

183. Horibata, K., Iwamoto, Y., Kuraoka, I., Jaspers, N.G., Kurimasa, A., Oshimura, M., Ichihashi, M. and Tanaka, K. (2004) Complete absence of Cockayne syndrome group B gene product gives rise to UV-sensitive syndrome but not Cockayne syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15410-15415.

184. Faghri, S., Tamura, D., Kraemer, K.H. and Digiovanna, J.J. (2008) Trichothiodystrophy: a systematic review of 112 published cases characterises a wide spectrum of clinical manifestations. *J Med Genet*, **45**, 609-621.

185. Giglia-Mari, G., Coin, F., Ranish, J.A., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N.G., Raams, A., Argentini, M., van der Spek, P.J. *et al.* (2004) A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat Genet*, **36**, 714-719.

186. Botta, E., Offman, J., Nardo, T., Ricotti, R., Zambruno, G., Sansone, D., Balestri, P., Raams, A., Kleijer, W.J., Jaspers, N.G. *et al.* (2007) Mutations in the C7orf11 (TTDN1) gene in six nonphotosensitive trichothiodystrophy patients: no obvious genotype-phenotype relationships. *Hum Mutat*, **28**, 92-96.

187. Nakabayashi, K., Amann, D., Ren, Y., Saarialho-Kere, U., Avidan, N., Gentles, S., MacDonald, J.R., Puffenberger, E.G., Christiano, A.M., Martinez-Mir, A. *et al.* (2005) Identification of C7orf11 (TTDN1) gene mutations and genetic heterogeneity in nonphotosensitive trichothiodystrophy. *Am J Hum Genet*, **76**, 510-516.

188. den Dulk, B., Sun, S.M., de Ruijter, M., Brandsma, J.A. and Brouwer, J. (2006) Rad33, a new factor involved in nucleotide excision repair in Saccharomyces cerevisiae. *DNA repair*, **5**, 683-692.

189. den Dulk, B., van Eijk, P., de Ruijter, M., Brandsma, J.A. and Brouwer, J. (2008) The NER protein Rad33 shows functional homology to human Centrin2 and is involved in modification of Rad4. *DNA repair*, **7**, 858-868.

190. Xie, Z., Liu, S., Zhang, Y. and Wang, Z. (2004) Roles of Rad23 protein in yeast nucleotide excision repair. *Nucleic acids research*, **32**, 5981-5990.

191. Ortolan, T.G., Chen, L., Tongaonkar, P. and Madura, K. (2004) Rad23 stabilizes Rad4 from degradation by the Ub/proteasome pathway. *Nucleic acids research*, **32**, 6490-6500.

192. Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K. and Weissman, J.S. (2003) Global analysis of protein expression in yeast. *Nature*, **425**, 737-741.

193. Jansen, L.E., Verhage, R.A. and Brouwer, J. (1998) Preferential binding of yeast Rad4.Rad23 complex to damaged DNA. *The Journal of biological chemistry*, **273**, 33111-33114.

194. Guzder, S.N., Sung, P., Prakash, L. and Prakash, S. (1998) Affinity of yeast nucleotide excision repair factor 2, consisting of the Rad4 and Rad23 proteins, for ultraviolet damaged DNA. *The Journal of biological chemistry*, **273**, 31541-31546.

195. Krasikova, Y.S., Rechkunova, N.I., Maltseva, E.A., Anarbaev, R.O., Pestryakov, P.E., Sugasawa, K., Min, J.H. and Lavrik, O.I. (2013) Human and yeast DNA damage recognition complexes bind with high affinity DNA structures mimicking in size transcription bubble. *Journal of molecular recognition : JMR*, **26**, 653-661.

196. Krasikova, Y.S., Rechkunova, N.I., Maltseva, E.A., Pestryakov, P.E., Petruseva, I.O., Sugasawa, K., Chen, X., Min, J.H. and Lavrik, O.I. (2013) Comparative analysis of interaction of human and yeast DNA damage recognition complexes with damaged DNA in nucleotide excision repair. *The Journal of biological chemistry*, **288**, 10936-10947.

197. Lafrance-Vanasse, J., Arseneault, G., Cappadocia, L., Legault, P. and Omichinski, J.G. (2013) Structural and functional evidence that Rad4 competes with Rad2 for binding to the Tfb1 subunit of TFIIH in NER. *Nucleic acids research*, **41**, 2736-2745.

198. Bardwell, A.J., Bardwell, L., Iyer, N., Svejstrup, J.Q., Feaver, W.J., Kornberg, R.D. and Friedberg, E.C. (1994) Yeast nucleotide excision repair proteins Rad2 and Rad4 interact with RNA polymerase II basal transcription factor b (TFIIH). *Molecular and cellular biology*, **14**, 3569-3576.

199. Chen, X., Velmurugu, Y., Zheng, G., Park, B., Shim, Y., Kim, Y., Liu, L., Van Houten, B., He, C., Ansari, A. *et al.* (2015) Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. *Nature communications*, **6**, 5849.

200. Velmurugu, Y., Chen, X., Slogoff Sevilla, P., Min, J.H. and Ansari, A. (2016) Twist-open mechanism of DNA damage recognition by the Rad4/XPC nucleotide excision repair complex. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E2296-2305.

201. Mu, H., Geacintov, N.E., Zhang, Y. and Broyde, S. (2015) Recognition of Damaged DNA for Nucleotide Excision Repair: A Correlated Motion Mechanism with a Mismatched cis-syn Thymine Dimer Lesion. *Biochemistry*, **54**, 5263-5267.

202. Boehr, D.D., Nussinov, R. and Wright, P.E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nature chemical biology*, **5**, 789-796.

203. Mu, H., Geacintov, N.E., Min, J.H., Zhang, Y. and Broyde, S. (2017) Nucleotide Excision Repair Lesion-Recognition Protein Rad4 Captures a Pre-Flipped Partner Base in a Benzo[a]pyrene-Derived DNA Lesion: How Structure Impacts the Binding Pathway. *Chemical research in toxicology*.

204. Brown, R. (1828) A brief account of microscopical observations made in the months of June, July and August1827,on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Edinburgh New Philosophical Journal* **5**, 358-371.

205. Brown, R. (1829) Additional remarks on active molecules. *Edinburgh Journal of Sciences* **1**, 314-320.

206. Kong, M. and 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.

207. Fick, A. (1855) Ueber Diffusion. *Annalen der Physik und Chemie*, **170**, 59-86.

208. Phillips, R., Kondev, J. and Theriot, J. (2009) *Physical biology of the cell*. Garland Science, New York.

209. Einstein, A. (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, **322**, 549-560.

210. Einstein, A. (1956) *Investigations on the Theory of the Brownian Movement*. Courier Corporation.

211. von Smoluchowski, M. (1906) Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Annalen der Physik*, **326**, 756-780.

212. Langevin, P. (1908) Sur la théorie du mouvement brownien. *C. R. Acad. Sci. (Paris)*, **146**, 530-533.

213. Pearson, K. (1905) The Problem of the Random Walk. *Nature*, **72**, 294-294.

214. Kalisky, T., Dekel, E. and Alon, U. (2007) Cost-benefit theory and optimal design of gene regulation functions. *Phys Biol*, **4**, 229-245.

215. Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F. *et al.* (1997) The complete genome sequence of Escherichia coli K-12. *Science*, **277**, 1453-1462.

216. Redding, S. and Greene, E.C. (2013) How do proteins locate specific targets in DNA? *Chemical physics letters*, **570**.

217. Lee, A.J., Warshaw, D.M. and Wallace, S.S. (2014) Insights into the glycosylase search for damage from single-molecule fluorescence microscopy. *DNA repair*, **20**, 23-31.

218. Tafvizi, A., Mirny, L.A. and van Oijen, A.M. (2011) Dancing on DNA: kinetic aspects of search processes on DNA. *Chemphyschem : a European journal of chemical physics and physical chemistry*, **12**, 1481-1489.

219. Kad, N.M., Wang, H., Kennedy, G.G., Warshaw, D.M. and 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**, 702-713.

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

221. Riggs, A.D., Bourgeois, S. and Cohn, M. (1970) The lac repressor-operator interaction. 3. Kinetic studies. *Journal of molecular biology*, **53**, 401-417.

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

223. Winter, R.B., Berg, O.G. and von Hippel, P.H. (1981) 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.

224. von Hippel, P.H. and Berg, O.G. (1989) Facilitated target location in biological systems. *The Journal of biological chemistry*, **264**, 675-678.

225. Halford, S.E. and Marko, J.F. (2004) How do site-specific DNA-binding proteins find their targets? *Nucleic acids research*, **32**, 3040-3052.

226. Slutsky, M. and Mirny, L.A. (2004) Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophysical journal*, **87**, 4021-4035.

227. Coppey, M., Benichou, O., Voituriez, R. and Moreau, M. (2004) Kinetics of target site localization of a protein on DNA: a stochastic approach. *Biophysical journal*, **87**, 1640-1649.

228. Loverdo, C., Benichou, O., Voituriez, R., Biebricher, A., Bonnet, I. and Desbiolles, P. (2009) Quantifying hopping and jumping in facilitated diffusion of DNA-binding proteins. *Physical review letters*, **102**, 188101.

229. Hu, T., Grosberg, A.Y. and Shklovskii, B.I. (2006) How proteins search for their specific sites on DNA: the role of DNA conformation. *Biophysical journal*, **90**, 2731-2744.

230. Lomholt, M.A., Ambjörnsson, T. and Metzler, R. (2005) Optimal Target Search on a Fast-Folding Polymer Chain with Volume Exchange. *Physical review letters*, **95**.

231. Lomholt, M.A., van den Broek, B., Kalisch, S.M., Wuite, G.J. and Metzler, R. (2009) Facilitated diffusion with DNA coiling. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 8204-8208.

232. Eliazar, I., Koren, T. and Klafter, J. (2007) Searching circular DNA strands. *Journal of Physics: Condensed Matter*, **19**, 065140.

233. Foffano, G., Marenduzzo, D. and Orlandini, E. (2012) Facilitated diffusion on confined DNA. *Physical Review E*, **85**.

234. Li, G.-W., Berg, O.G. and Elf, J. (2009) Effects of macromolecular crowding and DNA looping on gene regulation kinetics. *Nature Physics*, **5**, 294-297.

235. Elf, J., Li, G.W. and Xie, X.S. (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, **316**, 1191-1194.

236. Cravens, S.L., Schonhoft, J.D., Rowland, M.M., Rodriguez, A.A., Anderson, B.G. and Stivers, J.T. (2015) Molecular crowding enhances facilitated diffusion of two human DNA glycosylases. *Nucleic acids research*, **43**, 4087-4097.

237. Krepel, D., Gomez, D., Klumpp, S. and Levy, Y. (2016) Mechanism of Facilitated Diffusion during a DNA Search in Crowded Environments. *The journal of physical chemistry. B*, **120**, 11113-11122.

238. Brackley, C.A., Cates, M.E. and Marenduzzo, D. (2013) Intracellular Facilitated Diffusion: Searchers, Crowders, and Blockers. *Physical review letters*, **111**.

239. Liu, L. and Luo, K. (2014) Molecular crowding effect on dynamics of DNA-binding proteins search for their targets. *The Journal of Chemical Physics*, **141**, 225102.

240. Gomez, D. and Klumpp, S. (2016) Facilitated diffusion in the presence of obstacles on the DNA. *Physical chemistry chemical physics : PCCP*, **18**, 11184-11192.

241. Bauer, M. and Metzler, R. (2013) In vivo facilitated diffusion model. *PloS one*, **8**, e53956.

242. Meroz, Y., Eliazar, I. and Klafter, J. (2009) Facilitated diffusion in a crowded environment: from kinetics to stochastics. *Journal of Physics A: Mathematical and Theoretical*, **42**, 434012.

243. Gowers, D.M., Wilson, G.G. and Halford, S.E. (2005) Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 15883-15888.

244. Bonnet, I., Biebricher, A., Porte, P.L., Loverdo, C., Benichou, O., Voituriez, R., Escude, C., Wende, W., Pingoud, A. and Desbiolles, P. (2008) Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. *Nucleic acids research*, **36**, 4118-4127.

245. Tafvizi, A., Huang, F., Leith, J.S., Fersht, A.R., Mirny, L.A. and van Oijen, A.M. (2008) Tumor suppressor p53 slides on DNA with low friction and high stability. *Biophysical journal*, **95**, L01-03.

246. Gorman, J., Wang, F., Redding, S., Plys, A.J., Fazio, T., Wind, S., Alani, E.E. and Greene, E.C. (2012) Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E3074-3083.

247. Blainey, P.C., van Oijen, A.M., Banerjee, A., Verdine, G.L. and Xie, X.S. (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 5752-5757.

248. Ruusala, T. and Crothers, D.M. (1992) Sliding and intermolecular transfer of the lac repressor: kinetic perturbation of a reaction intermediate by a distant DNA sequence. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 4903-4907.

249. Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E.G., Berg, O.G. and Elf, J. (2012) The lac repressor displays facilitated diffusion in living cells. *Science*, **336**, 1595-1598.

250. Normanno, D., Boudarene, L., Dugast-Darzacq, C., Chen, J., Richter, C., Proux, F., Benichou, O., Voituriez, R., Darzacq, X. and Dahan, M. (2015) Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nature communications*, **6**, 7357.

251. Mirny, L.A., Slutsky, M., Wunderlich, Z., Tafvizi, A., Leith, J.S. and Kosmrlj, A. (2009) How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *Journal of Physics A: Mathematical and Theoretical*, **42**, 434013.

252. Kalodimos, C.G., Biris, N., Bonvin, A.M., Levandoski, M.M., Guennuegues, M., Boelens, R. and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science*, **305**, 386-389.

253. Hu, T. and Shklovskii, B.I. (2006) How does a protein search for the specific site on DNA: The role of disorder. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **74**, 021903.

254. Hu, L., Grosberg, A.Y. and Bruinsma, R. (2008) Are DNA transcription factor proteins maxwellian demons? *Biophysical journal*, **95**, 1151-1156.

255. Yu, S., Wang, S. and Larson, R.G. (2013) Proteins searching for their target on DNA by one-dimensional diffusion: overcoming the "speed-stability" paradox. *Journal of biological physics*, **39**, 565-586.

256. Bénichou, O., Kafri, Y., Sheinman, M. and Voituriez, R. (2009) Searching Fast for a Target on DNA without Falling to Traps. *Physical review letters*, **103**.

257. Reingruber, J. and Holcman, D. (2011) Transcription factor search for a DNA promoter in a three-state model. *Physical Review E*, **84**.

258. Bauer, M. and Metzler, R. (2012) Generalized facilitated diffusion model for DNA-binding proteins with search and recognition states. *Biophysical journal*, **102**, 2321-2330.

259. Tafvizi, A., Huang, F., Fersht, A.R., Mirny, L.A. and 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**, 563-568.

260. Leith, J.S., Tafvizi, A., Huang, F., Uspal, W.E., Doyle, P.S., Fersht, A.R., Mirny, L.A. and van Oijen, A.M. (2012) Sequence-dependent sliding kinetics of p53. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 16552-16557.

261. Gorman, J., Chowdhury, A., Surtees, J.A., Shimada, J., Reichman, D.R., Alani, E. and Greene, E.C. (2007) Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Molecular cell*, **28**, 359-370.

262. Cuculis, L., Abil, Z., Zhao, H. and Schroeder, C.M. (2015) Direct observation of TALE protein dynamics reveals a two-state search mechanism. *Nature communications*, **6**, 7277.

263. Hoogstraten, D., Bergink, S., Ng, J.M., Verbiest, V.H., Luijsterburg, M.S., Geverts, B., Raams, A., Dinant, C., Hoeijmakers, J.H., Vermeulen, W. *et al.* (2008) Versatile DNA damage detection by the global genome nucleotide excision repair protein XPC. *Journal of cell science*, **121**, 2850-2859.

264. Camenisch, U., Trautlein, D., Clement, F.C., Fei, J., Leitenstorfer, A., Ferrando-May, E. and Naegeli, H. (2009) Two-stage dynamic DNA quality check by xeroderma pigmentosum group C protein. *The EMBO journal*, **28**, 2387-2399.

265. Furda, A.M., Bess, A.S., Meyer, J.N. and 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.

266. Hughes, C.D., Wang, H., Ghodke, H., Simons, M., Towheed, A., Peng, Y., Van Houten, B. and Kad, N.M. (2013) Real-time single-molecule imaging reveals a direct interaction between UvrC and UvrB on DNA tightropes. *Nucleic acids research*, **41**, 4901-4912.

267. Parikh, D., Fouquerel, E., Murphy, C.T., Wang, H. and Opresko, P.L. (2015) Telomeres are partly shielded from ultraviolet-induced damage and proficient for nucleotide excision repair of photoproducts. *Nature communications*, **6**, 8214.

268. Pagano, J.M., Clingman, C.C. and Ryder, S.P. (2011) Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA*, **17**, 14-20.

269. Ortega, A., Amoros, D. and Garcia de la Torre, J. (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophysical journal*, **101**, 892-898.

270. Burchard, W., Schmidt, M. and Stockmayer, W.H. (1980) Information on Polydispersity and Branching from Combined Quasi-Elastic and Intergrated Scattering. *Macromolecules*, **13**, 1265-1272.

271. Arnspang, E.C., Brewer, J.R. and Lagerholm, B.C. (2012) Multi-color single particle tracking with quantum dots. *PloS one*, **7**, e48521.

272. Bagchi, B., Blainey, P.C. and Xie, X.S. (2008) Diffusion constant of a nonspecifically bound protein undergoing curvilinear motion along DNA. *The journal of physical chemistry. B*, **112**, 6282-6284.

273. Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M. *et al.* (1996) Life with 6000 genes. *Science*, **274**, 546, 563-547.

274. Hughes, B.D. (1995) *Random walks and random environments*. Clarendon Press; Oxford University Press, Oxford; New York.

275. Laughery, M.F., Hunter, T., Brown, A., Hoopes, J., Ostbye, T., Shumaker, T. and Wyrick, J.J. (2015) New vectors for simple and streamlined CRISPR-Cas9 genome editing in Saccharomyces cerevisiae. *Yeast*.

276. Bespalov, V.A., Conconi, A., Zhang, X., Fahy, D. and Smerdon, M.J. (2001) Improved method for measuring the ensemble average of strand breaks in genomic DNA. *Environ Mol Mutagen*, **38**, 166-174.

277. Janicijevic, A., Sugasawa, K., Shimizu, Y., Hanaoka, F., Wijgers, N., Djurica, M., Hoeijmakers, J.H. and Wyman, C. (2003) DNA bending by the human damage recognition complex XPC-HR23B. *DNA repair*, **2**, 325-336.

278. Yeo, J.E., Khoo, A., Fagbemi, A.F. and Scharer, O.D. (2012) The efficiencies of damage recognition and excision correlate with duplex destabilization induced by acetylaminofluorene adducts in human nucleotide excision repair. *Chemical research in toxicology*, **25**, 2462-2468.

279. Kong, M., Liu, L., Chen, X., Driscoll, K.I., Mao, P., Bohm, S., Kad, N.M., Watkins, S.C., Bernstein, K.A., Wyrick, J.J. *et al.* (2016) Single-Molecule Imaging Reveals that Rad4 Employs a Dynamic DNA Damage Recognition Process. *Molecular cell*, **64**, 376-387.

280. Verhage, R., Zeeman, A.M., de Groot, N., Gleig, F., Bang, D.D., van de Putte, P. and Brouwer, J. (1994) The RAD7 and RAD16 genes, which are essential for pyrimidine dimer removal from the silent mating type loci, are also required for repair of the nontranscribed strand of an active gene in Saccharomyces cerevisiae. *Molecular and cellular biology*, **14**, 6135-6142.

281. Kad, N.M. and Van Houten, B. (2012) Dynamics of lesion processing by bacterial nucleotide excision repair proteins. *Progress in molecular biology and translational science*, **110**, 1-24.

282. Komazin-Meredith, G., Mirchev, R., Golan, D.E., van Oijen, A.M. and Coen, D.M. (2008) Hopping of a processivity factor on DNA revealed by single-molecule assays of diffusion. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 10721-10726.

283. Friedberg, E.C., Friedberg, E.C. and American Society for Microbiology. (2006). 2nd ed. ASM Press,, Washington, D.C., pp. xxix, 1118 p.

284. Jaciuk, M., Nowak, E., Skowronek, K., Tanska, A. and Nowotny, M. (2011) Structure of UvrA nucleotide excision repair protein in complex with modified DNA. *Nature structural & molecular biology*, **18**, 191-197.

285. Luger, K. (2003) Structure and dynamic behavior of nucleosomes. *Curr Opin Genet Dev*, **13**, 127-135.

286. Nag, R. and Smerdon, M.J. (2009) Altering the chromatin landscape for nucleotide excision repair. *Mutation research*, **682**, 13-20.

287. Su, T.J., Tock, M.R., Egelhaaf, S.U., Poon, W.C. and Dryden, D.T. (2005) DNA bending by M.EcoKI methyltransferase is coupled to nucleotide flipping. *Nucleic acids research*, **33**, 3235-3244.

288. Mitchell, D.L., Haipek, C.A. and Clarkson, J.M. (1985) (6-4)Photoproducts are removed from the DNA of UV-irradiated mammalian cells more efficiently than cyclobutane pyrimidine dimers. *Mutation research*, **143**, 109-112.

289. Guzder, S.N., Sung, P., Prakash, L. and Prakash, S. (1997) Yeast Rad7-Rad16 complex, specific for the nucleotide excision repair of the nontranscribed DNA strand, is an ATP-dependent DNA damage sensor. *The Journal of biological chemistry*, **272**, 21665-21668.

290. Lettieri, T., Kraehenbuehl, R., Capiaghi, C., Livingstone-Zatchej, M. and Thoma, F. (2008) Functionally distinct nucleosome-free regions in yeast require Rad7 and Rad16 for nucleotide excision repair. *DNA repair*, **7**, 734-743.

291. Dunn, A.R., Kad, N.M., Nelson, S.R., Warshaw, D.M. and 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**, 7487-7498.

292. Lin, J., Countryman, P., Buncher, N., Kaur, P., E, L., Zhang, Y., Gibson, G., You, C., Watkins, S.C., Piehler, J. *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**, 2493-2504.

293. Hofling, F. and Franosch, T. (2013) Anomalous transport in the crowded world of biological cells. *Rep Prog Phys*, **76**, 046602.

294. Finkelstein, I.J. and Greene, E.C. (2013) Molecular traffic jams on DNA. *Annual review of biophysics*, **42**, 241-263.

295. Savir, Y. and Tlusty, T. (2007) Conformational proofreading: the impact of conformational changes on the specificity of molecular recognition. *PloS one*, **2**, e468.

296. Martin, D.S., Forstner, M.B. and Kas, J.A. (2002) Apparent subdiffusion inherent to single particle tracking. *Biophysical journal*, **83**, 2109-2117.

297. Liu, L.F. and Wang, J.C. (1987) Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 7024-7027.

298. Barbi, M., Place, C., Popkov, V. and Salerno, M. (2004) Base-sequence-dependent sliding of proteins on DNA. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **70**, 041901.

299. Barbi, M., Place, C., Popkov, V. and Salerno, M. (2004) A model of sequence-dependent protein diffusion along DNA. *Journal of biological physics*, **30**, 203-226.

300. Saxton, M.J. (2007) A Biological Interpretation of Transient Anomalous Subdiffusion. I. Qualitative Model. *Biophysical journal*, **92**, 1178-1191.

301. Zwanzig, R. (1988) Diffusion in a rough potential. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 2029-2030.

302. Lässig, M. (2007) From biophysics to evolutionary genetics: statistical aspects of gene regulation. *BMC bioinformatics*, **8**, S7.

303. Erie, D.A., Yang, G., Schultz, H.C. and Bustamante, C. (1994) DNA bending by Cro protein in specific and nonspecific complexes: implications for protein site recognition and specificity. *Science*, **266**, 1562-1566.

304. Slutsky, M., Kardar, M. and Mirny, L.A. (2004) Diffusion in correlated random potentials, with applications to DNA. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **69**, 061903.

305. Goychuk, I. and Kharchenko, V.O. (2014) Anomalous features of diffusion in corrugated potentials with spatial correlations: faster than normal, and other surprises. *Physical review letters*, **113**, 100601.

306. Lin, J., Countryman, P., Chen, H., Pan, H., Fan, Y., Jiang, Y., Kaur, P., Miao, W., Gurgel, G., You, C. *et al.* (2016) Functional interplay between SA1 and TRF1 in telomeric DNA binding and DNA-DNA pairing. *Nucleic acids research*, **44**, 6363-6376.

307. Lee, J.Y., Finkelstein, I.J., Arciszewska, L.K., Sherratt, D.J. and Greene, E.C. (2014) Single-molecule imaging of FtsK translocation reveals mechanistic features of protein-protein collisions on DNA. *Molecular cell*, **54**, 832-843.

308. Lomholt, M.A., Zaid, I.M. and Metzler, R. (2007) Subdiffusion and Weak Ergodicity Breaking in the Presence of a Reactive Boundary. *Physical review letters*, **98**.

309. D'Errico, M., Parlanti, E., Teson, M., de Jesus, B.M., Degan, P., Calcagnile, A., Jaruga, P., Bjoras, M., Crescenzi, M., Pedrini, A.M. *et al.* (2006) New functions of XPC in the protection of human skin cells from oxidative damage. *The EMBO journal*, **25**, 4305-4315.

310. Melis, J.P., Luijten, M., Mullenders, L.H. and van Steeg, H. (2011) The role of XPC: implications in cancer and oxidative DNA damage. *Mutation research*, **728**, 107-117.

311. Menoni, H., Hoeijmakers, J.H. and Vermeulen, W. (2012) Nucleotide excision repair-initiating proteins bind to oxidative DNA lesions in vivo. *The Journal of cell biology*, **199**, 1037-1046.

312. Yasuda, G., Nishi, R., Watanabe, E., Mori, T., Iwai, S., Orioli, D., Stefanini, M., Hanaoka, F. and Sugasawa, K. (2007) In vivo destabilization and functional defects of the xeroderma pigmentosum C protein caused by a pathogenic missense mutation. *Molecular and cellular biology*, **27**, 6606-6614.

313. Maillard, O., Solyom, S. and Naegeli, H. (2007) An aromatic sensor with aversion to damaged strands confers versatility to DNA repair. *PLoS biology*, **5**, e79.

314. Clement, F.C., Kaczmarek, N., Mathieu, N., Tomas, M., Leitenstorfer, A., Ferrando-May, E. and Naegeli, H. (2011) Dissection of the xeroderma pigmentosum group C protein function by site-directed mutagenesis. *Antioxidants & redox signaling*, **14**, 2479-2490.

315. Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y. *et al.* (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, **316**, 1160-1166.

316. Nguyen, T.A., Slattery, S.D., Moon, S.H., Darlington, Y.F., Lu, X. and Donehower, L.A. (2010) The oncogenic phosphatase WIP1 negatively regulates nucleotide excision repair. *DNA repair*, **9**, 813-823.

317. Reardon, J.T. and Sancar, A. (2003) Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes & development*, **17**, 2539-2551.

318. Hara, R., Mo, J. and Sancar, A. (2000) DNA damage in the nucleosome core is refractory to repair by human excision nuclease. *Molecular and cellular biology*, **20**, 9173-9181.

319. Horikoshi, N., Tachiwana, H., Kagawa, W., Osakabe, A., Matsumoto, S., Iwai, S., Sugasawa, K. and Kurumizaka, H. (2016) Crystal structure of the nucleosome containing ultraviolet light-induced cyclobutane pyrimidine dimer. *Biochemical and biophysical research communications*, **471**, 117-122.

320. Hara, R. and Sancar, A. (2002) The SWI/SNF chromatin-remodeling factor stimulates repair by human excision nuclease in the mononucleosome core particle. *Molecular and cellular biology*, **22**, 6779-6787.

321. Ura, K., Araki, M., Saeki, H., Masutani, C., Ito, T., Iwai, S., Mizukoshi, T., Kaneda, Y. and Hanaoka, F. (2001) ATP-dependent chromatin remodeling facilitates nucleotide excision repair of UV-induced DNA lesions in synthetic dinucleosomes. *The EMBO journal*, **20**, 2004-2014.

322. Marteijn, J.A., Lans, H., Vermeulen, W. and Hoeijmakers, J.H. (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nature reviews. Molecular cell biology*, **15**, 465-481.

323. Metzler, R., Jeon, J.H., Cherstvy, A.G. and Barkai, E. (2014) Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Physical chemistry chemical physics : PCCP*, **16**, 24128-24164.

324. Montroll, E.W. and Weiss, G.H. (1965) Random Walks on Lattices. II. *Journal of Mathematical Physics*, **6**, 167.

325. Scher, H. and Montroll, E.W. (1975) Anomalous transit-time dispersion in amorphous solids. *Physical Review B*, **12**, 2455-2477.

326. Metzler, R. and Klafter, J. (2004) The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *Journal of Physics A: Mathematical and General*, **37**, R161-R208.

327. Barkai, E., Garini, Y. and Metzler, R. (2012) Strange kinetics of single molecules in living cells. *Physics Today*, **65**, 29.

328. He, Y., Burov, S., Metzler, R. and Barkai, E. (2008) Random time-scale invariant diffusion and transport coefficients. *Physical review letters*, **101**, 058101.

329. Lubelski, A., Sokolov, I.M. and Klafter, J. (2008) Nonergodicity mimics inhomogeneity in single particle tracking. *Physical review letters*, **100**, 250602.

330. Weigel, A.V., Simon, B., Tamkun, M.M. and Krapf, D. (2011) Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 6438-6443.

331. Golding, I. and Cox, E.C. (2006) Physical nature of bacterial cytoplasm. *Physical review letters*, **96**, 098102.

332. Jeon, J.H., Tejedor, V., Burov, S., Barkai, E., Selhuber-Unkel, C., Berg-Sorensen, K., Oddershede, L. and Metzler, R. (2011) In vivo anomalous diffusion and weak ergodicity breaking of lipid granules. *Physical review letters*, **106**, 048103.

333. Mandelbrot, B.B. and Van Ness, J.W. (1968) Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review*, **10**, 422-437.

334. Chandler, D. (1987) *Introduction to modern statistical mechanics*. Oxford University Press, New York.

335. Lutz, E. (2001) Fractional Langevin equation. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **64**, 051106.

336. Kou, S.C. and Xie, X.S. (2004) Generalized Langevin equation with fractional Gaussian noise: subdiffusion within a single protein molecule. *Physical review letters*, **93**, 180603.

337. Jeon, J.H. and Metzler, R. (2010) Fractional Brownian motion and motion governed by the fractional Langevin equation in confined geometries. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **81**, 021103.

338. Deng, W. and Barkai, E. (2009) Ergodic properties of fractional Brownian-Langevin motion. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **79**, 011112.

339. Weiss, M., Elsner, M., Kartberg, F. and Nilsson, T. (2004) Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophysical journal*, **87**, 3518-3524.

340. Guigas, G., Kalla, C. and Weiss, M. (2007) The degree of macromolecular crowding in the cytoplasm and nucleoplasm of mammalian cells is conserved. *FEBS Lett*, **581**, 5094-5098.

341. Szymanski, J. and Weiss, M. (2009) Elucidating the origin of anomalous diffusion in crowded fluids. *Physical review letters*, **103**, 038102.

342. Ernst, D., Hellmann, M., Köhler, J. and Weiss, M. (2012) Fractional Brownian motion in crowded fluids. *Soft Matter*, **8**, 4886.

343. Weiss, M. (2013) Single-particle tracking data reveal anticorrelated fractional Brownian motion in crowded fluids. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **88**, 010101.

344. Magdziarz, M., Weron, A., Burnecki, K. and Klafter, J. (2009) Fractional brownian motion versus the continuous-time random walk: a simple test for subdiffusive dynamics. *Physical review letters*, **103**, 180602.

345. Weber, S.C., Spakowitz, A.J. and Theriot, J.A. (2010) Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm. *Physical review letters*, **104**, 238102.

346. Bronstein, I., Israel, Y., Kepten, E., Mai, S., Shav-Tal, Y., Barkai, E. and Garini, Y. (2009) Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Physical review letters*, **103**, 018102.

347. Kepten, E., Bronshtein, I. and Garini, Y. (2011) Ergodicity convergence test suggests telomere motion obeys fractional dynamics. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **83**, 041919.

348. Ben-Avraham, D. and Havlin, S. (2000) *Diffusion and reactions in fractals and disordered systems*. Cambridge University Press, Cambridge ; New York.

349. Saxton, M.J. (1994) Anomalous diffusion due to obstacles: a Monte Carlo study. *Biophysical journal*, **66**, 394-401.

350. Hellmann, M., Heermann, D.W. and Weiss, M. (2012) Enhancing phosphorylation cascades by anomalous diffusion. *EPL (Europhysics Letters)*, **97**, 58004.

351. Aoki, K., Yamada, M., Kunida, K., Yasuda, S. and Matsuda, M. (2011) Processive phosphorylation of ERK MAP kinase in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 12675-12680.

352. Saxton, M.J. (2001) Anomalous subdiffusion in fluorescence photobleaching recovery: a Monte Carlo study. *Biophysical journal*, **81**, 2226-2240.

353. Jeon, J.-H., Chechkin, A.V. and Metzler, R. (2014) Scaled Brownian motion: a paradoxical process with a time dependent diffusivity for the description of anomalous diffusion. *Phys. Chem. Chem. Phys.*, **16**, 15811-15817.

354. Cherstvy, A.G., Chechkin, A.V. and Metzler, R. (2013) Anomalous diffusion and ergodicity breaking in heterogeneous diffusion processes. *New Journal of Physics*, **15**, 083039.

355. Cherstvy, A.G. and Metzler, R. (2014) Nonergodicity, fluctuations, and criticality in heterogeneous diffusion processes. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **90**, 012134.

356. Kuhn, T., Ihalainen, T.O., Hyvaluoma, J., Dross, N., Willman, S.F., Langowski, J., Vihinen-Ranta, M. and Timonen, J. (2011) Protein diffusion in mammalian cell cytoplasm. *PloS one*, **6**, e22962.

357. Massignan, P., Manzo, C., Torreno-Pina, J.A., García-Parajo, M.F., Lewenstein, M. and Lapeyre, G.J. (2014) Nonergodic Subdiffusion from Brownian Motion in an Inhomogeneous Medium. *Physical review letters*, **112**.

358. Manzo, C., Torreno-Pina, J.A., Massignan, P., Lapeyre, G.J., Lewenstein, M. and Garcia Parajo, M.F. (2015) Weak Ergodicity Breaking of Receptor Motion in Living Cells Stemming from Random Diffusivity. *Physical Review X*, **5**.

359. Blumen, A., Klafter, J., White, B.S. and Zumofen, G. (1984) Continuous-Time Random Walks on Fractals. *Physical review letters*, **53**, 1301-1304.

360. Klafter, J., Blumen, A. and Zumofen, G. (1984) Fractal behavior in trapping and reaction: A random walk study. *Journal of Statistical Physics*, **36**, 561-577.

361. Tabei, S.M., Burov, S., Kim, H.Y., Kuznetsov, A., Huynh, T., Jureller, J., Philipson, L.H., Dinner, A.R. and Scherer, N.F. (2013) Intracellular transport of insulin granules is a subordinated random walk. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 4911-4916.

362. Meroz, Y., Sokolov, I.M. and Klafter, J. (2010) Subdiffusion of mixed origins: When ergodicity and nonergodicity coexist. *Physical Review E*, **81**.